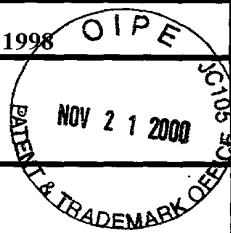


11414 Rec'd PCT/PTO 21 NOV 2000

FORM PTO-1390 (Modified) (REV 11-98)		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY'S DOCKET NUMBER 6433/80968 <i>Bot Selg</i>	
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371				U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR) 09/701132	
INTERNATIONAL APPLICATION NO. PCT/AU99/00385		INTERNATIONAL FILING DATE 21 MAY 1999		PRIORITY DATE CLAIMED 21 MAY 1998	
TITLE OF INVENTION ANTIGENS AND THEIR DETECTION					
APPLICANT(S) FOR DO/EO/US Peter Richard Reeves and Lei Wang					
Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:					
<ol style="list-style-type: none">1. <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.2. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.3. <input type="checkbox"/> This is an express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).4. <input type="checkbox"/> A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371 (c) (2))<ol style="list-style-type: none">a. <input checked="" type="checkbox"/> is transmitted herewith (required only if not transmitted by the International Bureau).b. <input checked="" type="checkbox"/> has been transmitted by the International Bureau.c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US).6. <input type="checkbox"/> A translation of the International Application into English (35 U.S.C. 371(c)(2)).7. <input type="checkbox"/> A copy of the International Search Report (PCT/ISA/210).8. <input checked="" type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))<ol style="list-style-type: none">a. <input checked="" type="checkbox"/> are transmitted herewith (required only if not transmitted by the International Bureau).b. <input checked="" type="checkbox"/> have been transmitted by the International Bureau.c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired.d. <input type="checkbox"/> have not been made and will not be made.9. <input type="checkbox"/> A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).10. <input type="checkbox"/> An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).11. <input type="checkbox"/> A copy of the International Preliminary Examination Report (PCT/IPEA/409).12. <input type="checkbox"/> A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)). <p>Items 13 to 20 below concern document(s) or information included:</p> <ol style="list-style-type: none">13. <input type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.14. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.15. <input checked="" type="checkbox"/> A FIRST preliminary amendment.16. <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment.17. <input type="checkbox"/> A substitute specification.18. <input type="checkbox"/> A change of power of attorney and/or address letter.19. <input type="checkbox"/> Certificate of Mailing by Express Mail20. <input type="checkbox"/> Other items or information: <div style="border: 1px solid black; height: 100px; width: 100%;"></div>					



U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR 1.53) 09/701132	INTERNATIONAL APPLICATION NO. PCT/AU99/00385	ATTORNEY'S DOCKET NUMBER 6433/80968
--	--	---

21. The following fees are submitted:.				CALCULATIONS PTO USE ONLY	
BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :					
<input type="checkbox"/> Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO \$1,000.00					
<input type="checkbox"/> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO \$860.00					
<input type="checkbox"/> International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO \$710.00					
<input type="checkbox"/> International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4) \$690.00					
<input type="checkbox"/> International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4) \$100.00					
ENTER APPROPRIATE BASIC FEE AMOUNT =				\$1,000.00	
Surcharge of \$130.00 for furnishing the oath or declaration later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492 (e)).				\$0.00	
CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	31 - 20 =	11	x \$18.00	\$198.00	
Independent claims	2 - 3 =	0	x \$80.00	\$0.00	
Multiple Dependent Claims (check if applicable). <input type="checkbox"/>				\$0.00	
TOTAL OF ABOVE CALCULATIONS =				\$1,198.00	
Reduction of 1/2 for filing by small entity, if applicable. Verified Small Entity Statement must also be filed (Note 37 CFR 1.9, 1.27, 1.28) (check if applicable). <input checked="" type="checkbox"/>				\$599.00	
SUBTOTAL =				\$599.00	
Processing fee of \$130.00 for furnishing the English translation later than <input type="checkbox"/> 20 <input type="checkbox"/> 30 months from the earliest claimed priority date (37 CFR 1.492 (f)).				\$0.00	
TOTAL NATIONAL FEE =				\$599.00	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) (check if applicable). <input type="checkbox"/>				\$0.00	
TOTAL FEES ENCLOSED =				\$599.00	
				Amount to be refunded	\$
				charged	\$

☒ A check in the amount of **\$599.00** to cover the above fees is enclosed.


☐ Please charge my Deposit Account No. _____ in the amount of _____ to cover the above fees.
A duplicate copy of this sheet is enclosed.

☒ The Commissioner is hereby authorized to charge any fees which may be required, or credit any overpayment to Deposit Account No. **23-0920** A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

--

SIGNATURE 

Shannon L. Nebolsky, Esq.

NAME

41,217

REGISTRATION NUMBER

November 20, 2000

DATE

528 Rec'd PCT/PTO 21 NOV 2000

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant: Peter Richard Reeves)
and Lei Wang) Attorney Docket:
6433/80968
U.S. Serial No.: Not yet assigned)
)
Filed: November 21, 2000)
)
Examining Group:
For: ANTIGENS AND THEIR)
DETECTION)
1600
)
Examiner: Not yet assigned)

PRELIMINARY AMENDMENT and SEQUENCE LISTING

Commissioner for Patents
Washington, D.C. 20231

Sir:

The subject application is a U.S. National Phase filing under 35 U.S.C. 371 based on International Patent Application Serial No. PCT/AU99/00385, international filing date May 21, 1999, claiming the benefit of foreign priority filing of Australian Patent Application Serial No. PP 3634, filed May 21, 1998.

ABSTRACT:

After the claims, please insert the following
Abstract of the invention.

--ABSTRACT

The invention relates to novel nucleotide sequences located in a gene which encodes a bacterial flagellin antigen, and the use of those nucleotide sequences for the detection of bacteria which express particular flagellin antigens, on the basis of that antigen alone, or in conjunction with the O antigen expressed by that strain.--

IN THE CLAIMS:

Please amend the claims as follows:

21. A method according to [any one of claims 8, 9, 11, 15 or 19] claim 7 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

22. A kit for identifying the H serotype of *E. coli*, the kit comprising at least one nucleic acid molecule according to [any one of claims 1 to 6] claim 1.

23. A kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule according to [any one of claims 1 to 6] claim 1; and

(b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or

processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

Please insert the following new claims.

26. A method according to claim 9 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

27. A method according to claim 11 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

28. A method according to claim 15 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

29. A method according to claim 19 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

30. A kit for identifying the H serotype of *E. coli*, the kit comprising at least one nucleic acid molecule according to claim 6.

31. A kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule according to claim 6; and

(b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

REMARKS

I. The Amendments

Claims 21 through 23 were amended to remove multiple dependency. New claims 26 to 31 were added to maintain the claimed subject matter as filed before removal of multiple dependency. The new claims are supported by the claims originally filed. A typographical error in claim dependency was also corrected in claim 21. No new matter has been added to the subject patent application by virtue of this preliminary amendments.

Claims 1 through 31 are in the case and are before the Examiner. An Abstract page has been added to the specification. The text for the abstract was taken from the PCT application abstract, thereby adding no new matter.

Pursuant to the new rules of practice in patent cases before the U.S. Patent and Trademark Office, a clean copy of the claims before the Examiner after entry of the Preliminary Amendment are enclosed. For the convenience of the Patent Office, a clean copy of the new Abstract page is also enclosed.

II. Biological Sequence Listing
Statements under 1.825(a) and 1.821(f)

This Preliminary Amendment including Statements under 37 C.F.R. 1.825(a) and 1.821(f) is accompanied by substitute sheets for the paper copy of the Sequence Listing of the above-identified patent application. The content of the Sequence Listing is the same as that of the Sequence Listing for the international application as filed in the PCT, the difference being that the format has been updated in the paper copy to conform to the current U.S. Patent Office Sequence Listing requirements with page numbers beginning at 1.

This paper is also accompanied by a write-protected diskette (3.50 inch, 1.44 Mb storage capacity) containing the computer readable form (CRF) of the Sequence Listing as ASCII output from PatentIn version 2.0. The computer readable form filename is

"P30384.app". The CRF of the sequence listing was generated by the PCT-filing associate in Australia using Patentin Version 2.0 on May 21, 1999 on a PC-compatible computer.

The Patentin output was transmitted via e-mail and copied onto the enclosed diskette November 21, 2000 unaltered as received. The information recorded in the computer readable form is identical to the enclosed paper copy of the Sequence Listing. A copy of the Patentin output was opened into a word processing program separately to produce the enclosed paper copy substitute sheets of the Biological Sequence Listing that has the appropriate page numbering. The substitute sheets include no new matter.

SUMMARY

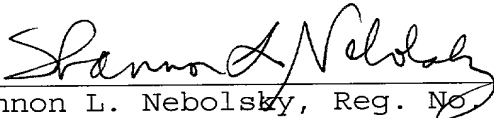
The claims and specification have been preliminarily amended to conform to U.S. practice, and substitute pages incorporating the preliminary amendments are enclosed, along with a computer-readable form of the Biological Sequence Listing to supplement the paper copy already transmitted from the international authority.

The application is believed to be in condition for allowance. An early notice to that effect is earnestly solicited.

No further fee or petition is believed to be necessary. However, should any further fee be needed, please charge our Deposit Account No. 23-0920, and deem this paper to be the required petition.

The Examiner is requested to phone the undersigned should any questions arise that can be dealt with over the phone to expedite this prosecution.

Respectfully submitted,


Shannon L. Nebolsky, Reg. No. 41,217

Enclosures:

Diskette with file P30384.app

Welsh & Katz, Ltd.
120 South Riverside Plaza
22nd Floor
Chicago, Illinois 60606
(312) 655-1500

CERTIFICATE OF EXPRESS MAILING

I hereby certify that this Preliminary Amendment and Sequence Listing, together with the stated enclosures, is being deposited with the United States Postal Service as Express Mail (EL617904151US) in an envelope addressed to: Commissioner for Patents, Washington, D.C. 20231 on November 21, 2000.

Slavko J. Petrov

CLAIMS

1. A nucleic acid molecule which encodes all or part of an *E. coli* flagellin protein, the molecule being capable of identifying the H serotype of an *E. coli* when hybridised to a gene of the *E. coli* which encodes a flagellin protein, provided that the molecule does not encode a flagellin protein expressed by the *E. coli* H1, H7, H12 or H48 type strains.
2. A nucleic acid molecule according to claim 1 wherein the molecule is derived from a *fliC* gene.
3. A nucleic acid molecule according to claim 1 including all or part of a sequence according to any one of SEQ ID NOs:1 to 68.
4. A nucleic acid molecule according to claim 1 consisting of all or part of a sequence according to any one of SEQ ID NOs: 1 to 68.
5. A nucleic acid molecule according to claim 4 wherein the molecule is from about 10 to 20 nucleotides in length.
6. A primer selected from the group of primers shown in Table 3.
7. A method of detecting the H serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1 in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and

(b) detecting a nucleic acid molecule which is hybridised to the gene, to detect the H serotype of the *E. coli* in the sample.

8. A method according to claim 7 wherein the hybridised nucleic acid molecules are detected by Southern Blot analysis.

9. A method of detecting the H serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of an *E. coli* in the sample with a pair of nucleic acid molecules according to claim 1 in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene; and

(b) detecting a pair of nucleic acid molecules which is hybridised to the gene, to detect the H serotype of the *E. coli* in the sample.

10. A method according to claim 9 wherein the hybridised pairs of nucleic acid molecules are detected by the polymerase chain reaction.

11. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of the *E. coli* with a nucleic acid molecule derived from a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a *E. coli* O antigen, in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene;

(b) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1 in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and

(c) detecting nucleic acid molecules which are hybridised to the genes, to detect the H and O serotype of the *E. coli* in the sample.

12. A method according to claim 11 wherein the nucleic acid molecule of step (a) is selected from the group consisting of:

wbdH (nucleotide position 739 to 1932 of Figure 5),
wzx (nucleotide position 8646 to 9911 of Figure 5),
wzy (nucleotide position 9901 to 10953 of Figure 5),
wbdM (nucleotide position 11821 to 12945 of Figure 5),
wbdN (nucleotide position 79 to 861 of Figure 6),
wbdO (nucleotide position 2011 to 2757 of Figure 6),
wbdP (nucleotide position 5257 to 6471 of Figure 6),
wbdR (nucleotide position 13156 to 13821 of Figure 6),
wzx (nucleotide position 2744 to 4135 of Figure 6) and
wzy (nucleotide position 858 to 2042 of Figure 6).

13. A method according to claim 12 wherein the nucleic acid molecule of step (a) is a primer

selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

14. A method according to claim 11 wherein the hybridised nucleic acid molecules are detected by Southern Blot analysis.

15. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of the *E. coli* with a pair of nucleic acid molecules derived from a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a *E. coli* O antigen, in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene;

(b) contacting a gene of an *E. coli* in the sample with a pair of nucleic acid molecules according to claim 1 in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene; and

(c) detecting pairs of nucleic acid molecules which are hybridised to the genes, to detect the H and O serotype of the *E. coli* in the sample.

16. A method according to claim 15 wherein the pair of nucleic acid molecules of step (a) is selected from the group consisting of:

wbdH (nucleotide position 739 to 1932 of Figure 5),

wzx (nucleotide position 8646 to 9911 of Figure 5),

wzy (nucleotide position 9901 to 10953 of Figure 5),
wbdM (nucleotide position 11821 to 12945 of Figure 5),
wbdN (nucleotide position 79 to 861 of Figure 6),
wbdO (nucleotide position 2011 to 2757 of Figure 6),
wbdP (nucleotide position 5257 to 6471 of Figure 6),
wbdR (nucleotide position 13156 to 13821 of Figure 6),
wzx (nucleotide position 2744 to 4135 of Figure 6) and
wzy (nucleotide position 858 to 2042 of Figure 6).

17. A method according to claim 15 wherein the nucleic acid molecules of the pair of step (a) are primers selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

18. A method according to claim 15 wherein the hybridised pairs of nucleic acid molecules are detected by the polymerase chain reaction.

19. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1, in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and

(b) detecting a nucleic acid molecule which is hybridised to the gene, to detect the H and O serotype of *E. coli* in the sample.

20. A method according to claim 19 wherein the nucleic acid molecule is according to any one of SEQ ID NOS: 9, 55, 57 to 65.

21. A method according to claim 7 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

22. A kit for identifying the H serotype of *E. coli*, the kit comprising at least one nucleic acid molecule according to claim 1.

23. A kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule according to claim 1; and

(b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

24. A kit according to claim 23 wherein the at least one nucleic acid molecule of (a) is selected from the group

consisting of:

wbdH (nucleotide position 739 to 1932 of Figure 5),
wzx (nucleotide position 8646 to 9911 of Figure 5),
wzy (nucleotide position 9901 to 10953 of Figure 5),
wbdM (nucleotide position 11821 to 12945 of Figure 5),
wbdN (nucleotide position 79 to 861 of Figure 6),
wbdO (nucleotide position 2011 to 2757 of Figure 6),
wbdP (nucleotide position 5257 to 6471 of Figure 6),

wbdR (nucleotide position 13156 to 13821 of Figure 6),
wzx (nucleotide position 2744 to 4135 of Figure 6) and
wzy (nucleotide position 858 to 2042 of Figure 6).

25. A kit according to claim 24 wherein the nucleic acid molecule of (a) is a primer selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

26. A method according to claim 9 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

27. A method according to claim 11 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

28. A method according to claim 15 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

29. A method according to claim 19 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

30. A kit for identifying the H serotype of *E. coli*, the kit comprising at least one nucleic acid molecule according to claim 6.

31. A kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule according to claim 6; and

(b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

ABSTRACT

The invention relates to novel nucleotide sequences located in a gene which encodes a bacterial flagellin antigen, and the use of those nucleotide sequences for the detection of bacteria which express particular flagellin antigens, on the basis of that antigen alone, or in conjunction with the O antigen expressed by that strain.

1. A nucleotide sequence located in a gene which encodes a bacterial flagellin antigen, and the use of those nucleotide sequences for the detection of bacteria which express particular flagellin antigens, on the basis of that antigen alone, or in conjunction with the O antigen expressed by that strain.

96/PRTS 09/701132
SECRET 21 NOV 2000

Antigens and Their Detection

TECHNICAL FIELD

The invention relates to novel nucleotide sequences located in a gene which encodes a bacterial flagellin antigen, and the use of those nucleotide sequences for the detection of bacteria which express particular flagellin antigens, on the basis of that antigen alone, or in conjunction with the O antigen expressed by that strain.

BACKGROUND ART

The flagellum of many bacteria appears to be made up of a single protein known as flagellin. The serotyping schemes of *E. coli* and *Salmonella enterica* are based on highly variable antigenic surface structures which include the lipopolysaccharide which carries the O antigen and flagellin which is now known to be the carrier of the classical H antigen. In many strains of *S. enterica* there are two loci (*fliC* and *fljB*) which encode flagellin, and a regulatory system which allows one only to be expressed at any time; and which also provides for expression to rapidly alternate between the two forms first identified as two phases (H1 and H2) for the H antigen of most strains. In *E. coli* there are 54 forms of H antigen recognised and until recently they were all thought to be encoded at the *fliC* locus, as has been shown for *E. coli* K-12. However in the 1980s Ratiner [Ratiner Y A "Phase variation of the H antigen in *Escherichia coli* strain Bi327-41, the standard strain for *Escherichia coli* flagellin antigen H3" FEMS Microbiol. Lett 15 (1982) 33-36; Ratiner Y A "Presence of two structural genes determining antigenically different phase-specific flagellins in some *Escherichia coli* strains" FEMS Microbiol. Lett. 19 (1983) 37-41; Ratiner Y A "Two genetic arrangements determining flagellin antigen specificities in two diphasic *Escherichia coli* strains" FEMS Microbiol. Lett. 29 (1985) 317-323; Ratiner Y A "Different alleles of the flagellin gene *hagB* in *Escherichia coli* standard H

test strains" FEMS Microbiol Lett. 48 (1987) 97-104.] showed that in some cases there are two loci and that expression can alternate. The matter was further complicated by a recent paper by Ratiner [Ratiner Y A (1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984] showing three loci (*flk*, *fll* and *flm*) for flagellin in addition to *fliC* although the *fljB* locus has not been found in *E. coli*. However *E. coli* strains are normally identified by the combination of one O antigen and one H antigen [and K antigen when present as a capsule (K) antigen], with no problems reported for the vast majority of cases with alternate phases, while *S. enterica* strains are normally identified by the combination of O, H1 and H2 antigens. It is still not clear how widespread in *E. coli* H antigens determined by flagellin genes other than *fliC* are.

Typing is typically carried out using specific antisera. The incidence of pathogenic *E. coli* in association with human and animal disease supports the need for suitable and rapid typing techniques.

DESCRIPTION OF THE INVENTION

In a first aspect, the present invention provides a novel nucleic acid molecule encoding all or part of an *E. coli* flagellin protein.

The present invention provides, for the first time, full length sequence for a flagellin gene for the following *E. coli* type strains: H6 (SEQ ID NO: 8), H9 (SEQ ID NO: 11), H10 (SEQ ID NO: 12), H14 (SEQ ID NO: 15), H18 (SEQ ID NO: 18), H23 (SEQ ID NO: 22), H51 (SEQ ID NO: 50), H45 (SEQ ID NO: 43), H49 (SEQ ID NO: 48), H19 (SEQ ID NO: 19), H30 (SEQ ID NO: 29), H32 (SEQ ID NO: 31), H26 (SEQ ID NO: 25), H41 (SEQ ID NO: 39), H15 (SEQ ID NO: 16), H20 (SEQ ID NO: 20), H28 (SEQ ID NO: 27), H46 (SEQ ID NO: 44), H31 (SEQ ID NO: 30), H34 (SEQ ID NO: 33), H43 (SEQ ID NO: 41) and H52 (SEQ ID NO: 51). Corrected full length sequences have been obtained for H7 (SEQ ID NO: 9) and

- 3 -

H12(SEQ ID NO: 14) type strains.

Partial flagellin gene sequence, including the central variable region, has been obtained for the following *E. coli* H type strains: H40(SEQ ID NO: 38), H8(SEQ ID NO: 10), H21(SEQ ID NO: 21), H47(SEQ ID NO: 46), H11(SEQ ID NO: 13), H17(SEQ ID NO: 17), H25(SEQ ID NO: 24), H42(SEQ ID NO: 40), H27(SEQ ID NO: 26), H35(SEQ ID NO: 34), H2(SEQ ID NO: 67), H3(SEQ ID NO: 68), H24(SEQ ID NO: 23), H37(SEQ ID NO: 35), H50(SEQ ID NO: 49), H4(SEQ ID NO: 6), H44(SEQ ID NO: 42), H38(SEQ ID NO: 36), H39(SEQ ID NO: 37), H55(SEQ ID NO: 53), H29(SEQ ID NO: 28), H33(SEQ ID NO: 32), H5(SEQ ID NO: 7), H54(SEQ ID NO: 52) and H56(SEQ ID NO: 54).

Comparison of sequences demonstrates that unique flagellin genes have now been sequenced (partially or completely) for the following *E. coli* H type strains: H1, H2, H3, H5, H6, H7, H9, H11, H12, H14, H15, H18, H19, H20, H21, H23, H24, H25, H26, H27, H28, H29, H30, H31, H32, H33, H34, H35, H37, H38, H39, H41, H42, H43, H45, H46, H48, H49, H51, H52, H54, and H56 and either H8 or H40, H10 or H50 and H4 or H17.

By comparison of these sequences, the present inventors were able to identify specific sequences for each of the above H serotypes.

The present invention also provides *fliC* sequences from 10 different H7 strains, in addition to that from the H7 type strain, and two sequences specific to H7 of O157 and O55 *E. coli* strains.

The present invention encompasses all or part of the flagellin genes sequenced for H2, H3, H5, H6, H9, H11, H14, H18, H19, H20, H21, H23, H24, H25, H26, H27, H28, H29, H30, H31, H32, H33, H34, H35, H37, H38, H39, H41, H42, H43, H44, H45, H46, H47, H48, H49, H51, H52, H54, H55, H56, H8, H40, H15, H10, or H50, H4 and H17 type strains. Of these flagellin genes sequenced, those from the type strains for H8 and H40 are identical, those from type strains H10 and H50, H1 and H12, H38 and H55, H21 and

H47, and H4, H17 and H44 type strains are highly similar.

The invention also encompasses newly provided sequence for H7 and H12 as well as novel primers for the specific amplification of H1, H7, H12 and H48 as well as for the other above mentioned newly sequenced flagellin genes.

By cloning and expression of these sequenced flagellin genes in a *fliC* deletion *E. coli* K-12 strain, and use of anti-H antiserum, we have confirmed the H specificities encoded by 39 flagellin genes. The 39 H specificities are H1, H2, H4, H5, H6, H7, H9, H10, H11, H12, H14, H15, H16, H18, H19, H20, H21, H23, H24, H26, H27, H28, H29, H30, H31, H32, H33, H34, H38, H39, H41, H42, H43, H45, H46, H49, H51, H52, and H56, encoded by flagellin genes obtained from H type strains for H1, H2, H4, H5, H6, H7, H9, H10, H11, H12, H14, H15, H18, H19, H20, H21, H23, H24, H26, H27, H28, H29, H30, H31, H32, H33, H34, H38, H39, H41, H42, H43, H45, H46, H49, H51, H52, and H56 respectively.

The nucleic acid molecules of the invention may be variable in length. In one embodiment they are oligonucleotides of from about 10 to about 20 nucleotides in length. The oligonucleotides of the invention are specific for the flagellin gene from which they are derived and are derived from the central region of the gene. In one embodiment, oligonucleotides in accordance with the present invention, which also include oligonucleotides from the previously sequenced *E. coli* H1, H7, H12 and H48 genes, are those shown in Table 3.

The 45 sequences (see Table 3) provide a panel to which newly sequenced genes can be compared to select specific oligonucleotides for those newly sequenced genes.

In a second aspect the invention provides a method of detecting the presence of *E. coli* of a particular H serotype in a sample, the method comprising the step of specifically hybridising at least one nucleic acid molecule derived from a flagellin gene, wherein the at

least one nucleic acid molecule is specific for a particular flagellin gene associated with the H serotype, to any *E. coli* in the sample which contain the gene, and detecting any specifically hybridised nucleic acid molecules, wherein the presence of specifically hybridised nucleic acid molecules identifies the presence of the H serotype in the sample.

In one preferred embodiment the detection method is a Southern blot method. More preferably, the nucleic acid molecule is labelled and hybridisation of the nucleic acid molecule is detected by autoradiography or detection of fluorescence.

Preferred nucleic acid molecules for the detection of particular flagellin genes are listed in Table 3.

In a third aspect the invention provides a method of detecting the presence of *E. coli* of a particular H serotype in a sample, the method comprising the step of specifically hybridising at least one pair of nucleic acid molecules to any *E. coli* in the sample which contains the flagellin gene for the particular H serotype, wherein at least one of the nucleic acid molecules is specific for the particular flagellin gene associated with the H serotype, and detecting any specifically hybridised nucleic acid molecules, wherein the presence of specifically hybridised nucleic acid molecules identifies the presence of the H serotype in the sample.

In one preferred embodiment the detection method is a polymerase chain reaction method. More preferably, the nucleic acid molecules are labelled and hybridisation of the nucleic acid molecule is detected by electrophoresis.

It is recognised that there may be instances where spurious hybridisation will arise through the initial selection of a sequence found in many different genes but this is typically recognisable by, for instance, comparison of band sizes against controls in PCR gels, and an alternative sequence can be selected.

In a fourth aspect the invention provides a method for detecting the presence of a particular O serotype and H serotype of *E. coli* in a sample, the method comprising the following steps:

5 (a) specifically hybridising at least one nucleic acid molecule, derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the
10 synthesis of a particular *E. coli* O antigen, to any *E. coli* in the sample which contain the gene;

(b) specifically hybridising at least one nucleic acid molecule derived from and specific for a particular flagellin gene associated with that H serotype, to any *E.*
15 *coli* in the sample which contain the gene; and

(c) detecting any specifically hybridised nucleic acid molecules.

Preferred nucleic acid molecules for the detection of particular flagellin genes are listed in Table 3.

20 In one preferred embodiment, the sequence of the nucleic acid molecule specific for the O antigen is specific to the nucleotide sequence encoding the O111 antigen. More preferably, the sequence is derived from a gene selected from the group consisting of *wbdH*
25 (nucleotide position 739 to 1932 of Figure 5), *wzx* (nucleotide position 8646 to 9911 of Figure 5), *wzy* (nucleotide position 9901 to 10953 of Figure 5), *wbdM* (nucleotide position 11821 to 12945 of Figure 5) and fragments of those molecules of at least 10-12 nucleotides
30 in length. Particularly preferred nucleic acid molecules are those set out in Tables 8 and 8A, with respect to the above mentioned genes.

In another preferred embodiment, the sequence of the nucleic acid molecule specific for the O antigen is
35 specific to the nucleotide sequence encoding the O157 antigen. More preferably, the sequence is derived from a gene selected from the group consisting of *wbdN*

(nucleotide position 79 to 861 of Figure 6), *wbdO* (nucleotide position 2011 to 2757 of Figure 6), *wbdP* (nucleotide position 5257 to 6471 of Figure 6), *wbdR* (nucleotide position 13156 to 13821 of Figure 6), *wzx* (nucleotide position 2744 to 4135 of Figure 6) and *wzy* (nucleotide position 858 to 2042 of Figure 6) and fragments of those molecules of at least 10-12 nucleotides in length. Particularly preferred nucleic acid molecules are those set out in Tables 9 and 9A, with respect to the above mentioned genes.

In one preferred embodiment the detection method is a Southern blot method. More preferably, the nucleic acid molecule is labelled and hybridisation of the nucleic acid molecule is detected by autoradiography or detection of fluorescence.

In a fifth aspect the invention provides a method for detecting the presence of a particular O serotype and H serotype of *E. coli* in a sample, the method comprising the following steps:

(a) specifically hybridising at least one pair of nucleic acid molecules, at least one of which is derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of the particular *E. coli* O antigen, to any *E. coli* in the sample which contain the gene;

(b) specifically hybridising at least one pair of nucleic acid molecules, at least one of which is derived from and specific for a particular flagellin gene associated with the particular H serotype, to any *E. coli* in the sample which contain the gene; and

(c) detecting any specifically hybridised nucleic acid molecules.

Preferred nucleic acid molecules for the detection of particular flagellin genes are listed in Table 3.

In one preferred embodiment, the sequence of the nucleic acid molecule specific for the O antigen is specific to the nucleotide sequence encoding the O111 antigen. More preferably, the sequence is derived from a gene selected from the group consisting of *wbdH* (nucleotide position 739 to 1932 of Figure 5), *wzx* (nucleotide position 8646 to 9911 of Figure 5), *wzy* (nucleotide position 9901 to 10953 of Figure 5), *wbdM* (nucleotide position 11821 to 12945 of Figure 5) and fragments of those molecules of at least 10-12 nucleotides in length. Particularly preferred nucleic acid molecules are those set out in Tables 8 and 8A, with respect to the above mentioned genes.

In another preferred embodiment, the sequence of the nucleic acid molecule specific for the O antigen is specific to the nucleotide sequence encoding the O157 antigen. More preferably, the sequence is derived from a gene selected from the group consisting of *wbdN* (nucleotide position 79 to 861 of Figure 6), *wbdO* (nucleotide position 2011 to 2757 of Figure 6), *wbdP* (nucleotide position 5257 to 6471 of Figure 6), *wbdR* (nucleotide position 13156 to 13821 of Figure 6), *wzx* (nucleotide position 2744 to 4135 of Figure 6) and *wzy* (nucleotide position 858 to 2042 of Figure 6) and fragments of those molecules of at least 10-12 nucleotides in length. Particularly preferred nucleic acid molecules are those set out in Tables 9 and 9A, with respect to the above mentioned genes.

In one preferred embodiment the detection method is a polymerase chain reaction method. More preferably, the nucleic acid molecules are labelled and hybridisation of the nucleic acid molecule is detected by electrophoresis.

The present inventors believe that based on the teachings of the present invention and available information concerning O antigen gene clusters, and through use of experimental analysis, comparison of nucleic acid sequences or predicted protein structures, nucleic acid molecules in accordance with the invention

can be readily derived for any particular O antigen of interest. Suitable bacterial strains can typically be acquired commercially from depositary institutions.

There are currently 166 defined *E. coli* O antigens.

5 Samples of the 166 different *E. coli* O antigen serotypes are available from Statens Serum Institut, Copenhagen, Denmark.

10 The inventors envisage rare circumstances whereby two genetically similar gene clusters encoding serologically different O antigens have arisen through recombination of genes or mutation so as to generate polymorphic variants.

15 In these circumstances multiple pairs of oligonucleotides may be selected to provide hybridisation to the specific combination of genes. The invention thus envisages the use of a panel containing multiple nucleic acid molecules for use in the method of testing for O antigen in conjunction with H antigen, wherein the nucleic acid molecules are derived from genes encoding transferases and/or enzymes for the transport or processing of a polysaccharide or oligosaccharide unit including *wzx* or *wzy* genes, wherein the panel of nucleic acid molecules is specific to a particular O antigen. The panel of nucleic acid molecules can include nucleic acid molecules derived from O antigen sugar pathway genes where necessary.

20 The inventors also found two mutated flagellin genes from H type strains for H35 and H54 which have insertion sequences inserted into normal flagellar genes identical or near identical to that that of the H11 and H21 type strains respectively. Thus, primers for H11 and H21
30 (listed in Table 3) would also amplify fragments in H35 and H54, which differ in sizes to those in H11 and H21 respectively. The inventors also provide two pairs of primers each for H35 and H54 based on the insertion sequence (see H35 and H54 columns in Table 3). The use of
35 one of them in combination with one of the H11 or H21 primers will generate a PCR band only in H35 or H54 respectively, and this will also differentiate H35 and H54

from H11 and H21 respectively.

The present invention also relates to methods of detecting the presence of particular *E. coli* H antigens or H antigen and O antigen combinations where one or more nucleic acid molecules which generate a particular size fragment indicative of the presence of that H antigen are used or in which the combination of one antigen specific primer for that H antigen with another primer for a related H antigen provides for the detection of the particular H antigen by hybridisation to the relevant gene. Preferably, the H antigen is H11, H21, H35 or H54.

The pairs of nucleic acid molecules where the method of the fifth aspect is used may both hybridise to the relevant H or O antigen gene or alternatively only one may hybridise to the relevant gene and the other to another site.

The inventors recognise in applying the methods of the invention for detecting combinations of O and H antigens to samples, that the methods do not indicate whether a positive result for a particular O and H antigen combination arises because the O and H antigen are present on a single *E. coli* strain present in the sample or are present on different *E. coli* strains present in the sample. Because the ability to identify the presence of *E. coli* strains with particular O and H antigen combinations is highly desirable (due to the relationship between particular combinations and pathogenicity) the determination that a particular combination is present in a sample can be followed by isolation of single colonies and checking whether they contain the relevant combination by using the same method again or using antibody labelled magnetic beads to separate cells expressing the particular O or H antigen and then testing the isolated cells for the other serotype.

In addition, as mentioned above, the present inventors have established the existence of H7 primers specific to the O157 and O55 serotypes. Using such

primers it is possible to detect particular O and H antigen combinations with the use of H specific nucleic acid molecules.

In a sixth aspect the invention provides a method for detecting the presence of a particular O serotype and H serotype of *E. coli* in a sample, the method comprising the following steps:

(a) specifically hybridising at least one nucleic acid molecule, derived from and specific for a gene encoding a flagellin associated with a particular *E. coli* H antigen serotype to any *E. coli* carrying the gene and present in the sample;

and

(b) detecting the at least one specifically hybridised nucleic acid molecule, wherein the at least one nucleic acid molecule is specific for the particular combination of O and H antigen.

Preferably the combination is O55:H7 or O157:H7.

The ability to detect the O157:H7 combination from a particular H7 primer or pair is of particular use given the association of this combination with pathogenic strains.

In a seventh aspect the present invention provides a method for testing a food derived sample for the presence of one or more particular *E. coli* O antigens and H antigens comprising testing the sample by a method of the fourth, fifth or sixth aspect the invention.

In an eighth aspect the present invention provides a method for testing a faecal derived sample for the presence of one or more particular *E. coli* O antigens and H antigens comprising testing the sample by a method of the fourth, fifth or sixth aspect the invention.

In a ninth aspect the present invention provides a method for testing a patient or animal derived sample for the presence of one or more particular *E. coli* O antigens and H antigens comprising testing the sample by a method of the fourth, fifth or sixth aspect the invention.

- 12 -

Preferably, the method of the seventh, eighth or ninth aspect of the invention is a polymerase chain reaction method. More preferably the oligonucleotide molecules for use in the method are labelled. Even more preferably the hybridised nucleic acid molecules are detected by electrophoresis.

In the above described methods it will be understood that where pairs of nucleic acid molecules are used one of the nucleic acid molecules may hybridise to a sequence that is not from the O antigen transferase, *wzx* or *wzy* gene or the flagellin gene. Further where both hybridise to these genes the O antigen molecules may hybridise to the same or a different one of these genes.

In a tenth aspect the present invention provides a kit for identifying the H serotype of *E. coli*, the kit comprising:

at least one nucleic acid molecule derived from and specific for an *E. coli* flagellin gene.

In an eleventh aspect the present invention provides a kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule derived from and specific for an *E. coli* flagellin gene; and

(b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

The nucleic acid molecules may be provided in the same or different vials. The kit may also provide in the same or separate vials a second set of specific nucleic acid molecules.

Particularly preferred nucleic acid molecules for inclusion in the kits are those specified in Tables 3, 8, 8A, 9 and 9A as described above.

DEFINITIONS

In this specification, we have used term "flagellin gene" in many cases where previously one would have used "*fliC*", to allow for the uncertainty as to locus introduced by recent observations. However, uncertainty as to the locus does not alter the fact that most *E. coli* strains express a single H antigen and that a single flagellin gene sequence per strain is required to give the genetic basis for H antigen variation. Any use of the name *fliC* in this specification where a different locus is later shown to be involved would not affect the validity of conclusions drawn regarding application of information based on the sequence, where the conclusions do not relate to the map position. Thus it is generally the nucleic acid molecule itself which is of importance rather than the name attributed to the gene. When it is known or suspected that the gene encoding the H antigen is not in the *fliC* locus, we use the term flagellin rather than *fliC*.

The phrase, "a nucleic acid molecule derived from a gene" means that the nucleic acid molecule has a nucleotide sequence which is either identical or substantially similar to all or part of the identified gene. Thus a nucleic acid molecule derived from a gene can be a molecule which is isolated from the identified gene by physical separation from that gene, or a molecule which is artificially synthesised and has a nucleotide sequence which is either identical to or substantially similar to all or part of the identified gene. While some workers consider only the DNA strand with the same sequence as the mRNA transcribed from the gene, here either strand is intended.

Transferase genes are regions of nucleic acid which have a nucleotide sequence which encodes gene products that transfer monomeric sugar units.

Flippase or *wzx* genes are regions of nucleic acid which have a nucleotide sequence which encodes a gene

- 14 -

product that flips oligosaccharide repeat units generally composed of three to six monomeric sugar units to the external surface of the membrane.

Polymerase or *wzy* genes are regions of nucleic acid which have a nucleotide sequence which encodes gene products that polymerise repeating oligosaccharide units generally composed of 3-6 monomeric sugar units.

The nucleotide sequences provided in this specification are described as anti-sense sequences. This term is used in the same manner as it is used in Glossary of Biochemistry and Molecular Biology Revised Edition, David M. Glick, 1997 Portland Press Ltd., London on page 11 where the term is described as referring to one of the two strands of double-stranded DNA usually that which has the same sequence as the mRNA. We use it to describe this strand which has the same sequence as the mRNA.

NOMENCLATURE

Synonyms for *E. coli* O111 *rfb*

Current names	Our names	Bastin et al. 1991
wbdH	orf1	
gmd	orf2	
wbdI	orf3	orf3.4*
manC	orf4	rfbM*
manB	orf5	rfbK*
wbdJ	orf6	orf6.7*
wbdK	orf7	orf7.7*
wzx	orf8	orf8.9 and rfbX*
wzy	orf9	
wbdL	orf10	
wbdM	orf11	

* Nomenclature according to Bastin D.A., et al. 1991 "Molecular cloning and expression in *Escherichia coli* K-12 of the *rfb* gene cluster determining the O antigen of an *E. coli* O111 strain". Mol. Microbiol. 5:9 2223-2231.

Other Synonyms

wzy	rfc
wzx	rfbX
rmlA	rfbA
rmlB	rfbB
rmlC	rfbC
rmlD	rfbD
glf	orf6*
wbbI	orf3#, orf8* of <i>E. coli</i> K-12

- 15 -

wbbJ orf2#, orf9* of E. coli K-12
wbbK orf1#, orf10* of E. coli K-12
wbbL orf5#, orf 11* of E. coli K-12
Nomenclature according to Yao, Z. And M. A. Valvano 1994.

5 "Genetic analysis of the O-specific lipopolysaccharide biosynthesis region (rfb) of Escherichia coli K-12 W3110: identification of genes the confer groups-specificity to Shigella flexneri serotypes Y and 4a". J. Bacteriol. 176: 4133-4143.

10 * Nomenclature according to Stevenson et al. 1994. "Structure of the O-antigen of E. coli K-12 and the sequence of its rfb gene cluster". J. Bacteriol 176: 4144-4156.

15 • The O antigen genes of many species were given rfb names (rfbA etc) and the O antigen gene cluster was often referred to as the rfb cluster. There are now new names for the rfb genes as shown in the table. Both terminologies have been used herein, depending on the source of the information.

20 In the claims that follow and in the summary of the invention, except where the context requires otherwise due to express language or necessary implication, the word "comprising" is used in the sense of "including", i.e. the features specified may be associated with further features in various embodiments of the invention.

25 BRIEF DESCRIPTION OF THE DRAWINGS

30 Figure 1 shows *Eco* R1 restriction maps of cosmid clones pPR1054, pPR1055, pPR1056, pPR1058, pPR1287 which are subclones of *E. coli* O111 O antigen gene cluster. The thickened line is the region common to all clones. Broken lines show segments that are non-contiguous on the chromosome. The deduced restriction map for *E. coli* strain M92 is shown above.

35 Figure 2 shows a restriction mapping analysis of *E. coli* O111 O antigen gene cluster within the cosmid clone pPR1058. Restriction enzymes are: (B: *Bam*H1; Bg: *Bgl*III, E: *Eco*R1; H: *Hind*III; K: *Kpn*I; P: *Pst*I; S: *Sal*I and X: *Xho*I. Plasmids pPR1230, pPR1231, and pPR1288 are deletion derivatives of pPR1058. Plasmids pPR 1237, pPR1238, pPR1239 and pPR1240 are in pUC19. Plasmids pPR1243, pPR1244, pPR1245, pPR1246 and pPR1248 are in pUC18, and pPR1292 is in pUC19. Plasmid pPR1270 is in

40

- 16 -

pT7T319U. Probes 1, 2 and 3 were isolated as internal fragments of pPR1246, pPR1243 and pPR1237 respectively. Dotted lines indicate that subclone DNA extends to the left of the map into attached vector.

5 Figure 3 shows the structure of *E. coli* O111 O antigen gene cluster.

Figure 4 shows the structure of *E. coli* O157 O antigen gene cluster.

10 Figure 5 shows the nucleotide sequence (SEQ ID NO: 45) of the *E. coli* O111 O antigen gene cluster. Note: (1) The first and last three bases of a gene are underlined and of *italic* respectively.; (2) The region which was previously sequenced by Bastin and Reeves 1995 "Sequence and analysis of the O antigen gene (*rfb*) cluster of *Escherichia coli* O111" Gene 164: 17-23 is marked.

15 Figure 6 shows the nucleotide sequence (SEQ ID NO: 56) of the *E. coli* O157 O antigen gene cluster. Note: (1) The first and last three bases of a gene (region) are underlined and of *italic* respectively (2) The region previously sequenced by Bilge et al. 1996 "Role of the *Escherichia coli* O157-H7 O side chain in adherence and analysis of an *rfb* locus". Inf. and Immun 64:4795-4801 is marked.

20 Figures 7 to 9 show the nucleotide sequences (SEQ ID NOS: 66 to 68 respectively) obtained for flagellin genes from *E. coli* type strains for H1 to H3 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

25 Figures 10 to 18 show the nucleotide sequences (SEQ ID NOS: 6 to 14 respectively) obtained for flagellin genes from *E. coli* type strains for H4 to H12 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

30 Figures 19 and 20 show the nucleotide sequences (SEQ ID NOS: 15 to 16 respectively) obtained for flagellin genes from *E. coli* type strains for H14 and H15 respectively. The primer positions listed in Table 3 are

based on treating the first nucleotide of each of these sequences as No. 1.

5 Figures 22 and 26 show the nucleotide sequences (SEQ ID NOS: 17 to 21 respectively) obtained for flagellin genes from *E. coli* type strains for H17 and H21 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

10 Figures 27 to 39 show the nucleotide sequences (SEQ ID NOS: 22 to 34) obtained for flagellin genes from *E. coli* type strains for H23 to H35 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

15 Figures 40 to 49 show the nucleotide sequences (SEQ ID NOS: 35 to 44) obtained for flagellin genes from *E. coli* type strains for H37 to H46 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

20 Figures 50 to 55 show the nucleotide sequences (SEQ ID NOS: 46 to 51) obtained for flagellin genes from *E. coli* type strains for H47 to H52 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

25 Figures 56 to 58 show the nucleotide sequences (SEQ ID NOS: 52 to 54) obtained for flagellin genes from *E. coli* type strains for H54 to H56 respectively. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

30 Figure 59 shows the nucleotide sequence (SEQ ID NO: 55) obtained for the flagellin gene from *E. coli* H7 strain M1179. The primer positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

35 Figures 60 to 68 show the nucleotide sequences (SEQ ID NOS: 57 to 65 respectively) obtained for flagellin genes from *E. coli* strains M1004, M1211, M1200, M1686, M1328, M917, M527, M973, and M918 respectively. The primer

positions listed in Table 3 are based on treating the first nucleotide of each of these sequences as No. 1.

Figure 69 shows the nucleotide sequence (SEQ ID NO: 1) of the *fliC* gene and DNA flanking the *fliC* gene from the H25 type strain.

Figure 70A shows the nucleotide sequence (SEQ ID NO: 2) obtained from the 5' end of the insert of plasmid pPR1989. The insert of plasmid pPR1989 encodes the second flagellin gene of the H55 type strain.

Figure 70B shows the nucleotide sequence (SEQ ID NO: 3) obtained from the 3' end of the insert of plasmid pPR1989. The insert of plasmid pPR1989 encodes the second flagellin gene of the H55 type strain.

Figure 71 shows the nucleotide sequence (SEQ ID NO:4) obtained from the 5' end of the insert of plasmid pPR1993. The insert of plasmid pPR1993 encodes the second flagellin gene of the H36 strain.

Figure 72 shows the nucleotide sequence (SEQ ID NO:5) obtained from the 3' end of the insert of plasmid pPR1993. The insert of plasmid pPR1993 encodes the second flagellin gene of the H36 type strain.

Figure 73 A shows the sequence of polylinker and the SD sequence of plasmid pTrc99A.

Figure 73B shows the sequence of the junction region between the SD sequence and the start of flagellin gene in the plasmids used for the expression of flagellin genes.

BEST METHOD OF CARRYING OUT THE INVENTION

In carrying out the methods of the invention with respect to the testing of particular sample types including samples from food, patients, animals and faeces the samples are prepared by routine techniques routinely used in the preparation of such samples for DNA based testing. The steps for testing the samples using particular nucleic acid molecules in assay formats such as Southern blots and PCR are performed under routinely determined conditions appropriate to the sample and the

nucleic acid molecules.

H antigen

Materials and Methods

5 1. Bacterial strains and plasmid:

There are 54 H types in *E. coli* [Ewing, W.H.: Edwards and Ewing's identification of the *Enterobacteriaceae.*, Elsevier Science Publishers, Amsterdam, The Netherlands, 1986]: note H antigens from 1 to 57 were listed and that 13, 22 and 57 are not valid. All the standard H type strains except H16 were obtained from the Institute of Medical and Veterinary Science, Adelaide, Australia. The primary stocks are held at the Statens Serum Institut, Copenhagen, Denmark.

15 The additional H7 strains used are listed in Table 1.

We do not have the type strain for H16. It is known that the H3 type strain is biphasic and can also express the H16 flagellin gene [Ratiner, Y. A. (1985) "Two genetic arrangements determining flagellar antigen specificities in two diphasic *E. coli* strains. FEMS Microbiol Lett 19: 317-323]. We have sequenced and cloned the H16 flagellin gene from the H3 type strain (see below).

20 *E. coli* K-12 strain C600 *hsm hsr fliC::Tn10* [Kuwajiwa, G. (1988) "Flagellin domain that affects H antigenicity of *E. coli* K-12" J. Bacteriol. 170; 485-488] (laboratory stock no. M2126) was obtained from Dr Benita Westerlund-Wikstrom of the Department of Biosciences, University of Helsinki, Finland. *E. coli* K-12 strain EJ2282 (laboratory no. P5560) is a *fliC* deletion strain, and was obtained from Dr Masatoshi Enomoto of the Department of Biology, Okayama University, Japan [Tominaga, A. M. A.-H. Mahmoud, T. Mokaiharu and M. Enomoto (1994) "Molecular characterization of intact but cryptic, flagellin genes in the genus *Shigella.*: Mol. Microbiol. 12: 277-285].

35 Plasmid pTrc99A was purchased from Pharmacia LKB (Melbourne, VIC, Australia).

2. Antisera

Antisera against H1, H3, H8, H14, H15, H17, H23, H24, H25, H26, H29, H30, H31, H32, H33, H35, H36, H37, H38, H39, H43, H44, H46, H47, H48, H49, H52, H53, H54, H55, and H56 were obtained from the Institute of Medical and veterinary Science, Adelaide, Australia. Antisera against H2, H4, H5, H6, H7, H9, H10, H11, H12, H16, H18, H19, H20, H21, H27, H28, H34, H40, H41, H42, H45, and H51 were obtained from Denka Seiken Co., Ltd, Tokyo, Japan.

Antisera to type H50 was not available from any known source.

The antisera available were checked against the appropriate type strains to confirm the specificities of both flagellin H antigen and H antisera: 52 sera (all those except anti-H16 serum listed above) gave a positive reaction with the corresponding type strains for that serum.

3. Agglutination test:

Bacteria from 1 ml of an overnight culture grown in Luria broth (Difco Tryptone, 10g/l; Difco yeast extract, 5g/l; NaCl, 0.5 g/l; pH 7.2) at 30°C was centrifuged (4000 rpm/10 min) and the bacteria pellet resuspended in 100 ml of saline. The agglutination test was carried out by mixing equal volumes (5 ml) of both the cells and antiserum on a slide. The slide was rocked for 1 minute and then observed for agglutination. For all agglutination tests, saline containing no antiserum was mixed with cells to be used as a negative control.

For testing the H specificities of strain M2126 or strain P5560 carrying plasmid containing cloned flagellin genes, cells of M2126 or P5560 were used as an additional negative control.

All agglutination tests were first carried out using undiluted antisera (note that the antisera we used have been diluted before reaching our hands), except for anti-

H11, anti-H34, anti-H52 and anti-H26 serum for which we used 1:10 dilutions to avoid background agglutination. In cases for which cross-reactions have been reported, we carried out agglutination tests using serial dilutions of sera (see section 10.1)

4. Motility test:

The motility of strain M2126 or strain P5560 carrying cloned flagellin genes was examined microscopically. 1 ml of overnight culture grown in Luria broth (Difco Tryptone, 10g/l; Difco yeast extract, 5g/l; NaCl, 0.5 g/l; pH 7.2) at 30°C was inoculated into 10 ml of Luria broth, and the culture was shaken at 100 rpm at 30°C to early log phase (OD 625 = 0.2). A loopful of culture was placed on a slide and examined under a microscope. Motility of individual cells was easily distinguished from Brownian movement and streaming, and presence or absence of motility recorded.

5. Isolation of chromosomal DNA:

Chromosomal DNA from all the 53 H type strains and the strains listed in Table 1 was isolated using the Promega Genomic isolation kit (Madison WI USA). Each chromosomal DNA sample was checked by gel electrophoresis of the DNA and by PCR amplification of the *mdh* gene using oligonucleotides based on the *E. coli* K-12 *mdh* gene [Boyd, E.F., Nelson, K., Wang, F.-S., Whittam, T.S. and Selander, R.K.: Molecular genetic basis of allelic polymorphism in malate dehydrogenase (*mdh*) in natural populations of *Escherichia coli* and *Salmonella enterica*. Proc. Natl. Acad. Sci. USA 91 (1994) 1280-1284].

6. PCR amplification of flagellin gene:

Flagellin genes from different strains were first PCR amplified using one of the following four pairs of oligonucleotides:

#1285 (5'-atggcacaagtcattaatac) and
#1286 (5'-ttaaccctgcagtagagaca);

- 22 -

#1417 (5'-ctgatcactcaaaataatatcaac) and
#1418 (5'-ctgcggtacctggttggc);
#1431 (5'-atggcacaagtcattaatacccaac) and
#1432 (5'-ctaaccctgcagcagagaca):
5 #1575 (5'-gggtggaaaccaatacg) and
#1576 (5'-gcgcatcaggcaatttgg)

PCR reactions were carried out under the following
conditions: denaturing, 94°C/30'; annealing, temperature
varies (refer to Table 2)/30'; extension, 72°C/1'; 30
10 cycles. The PCR product was purified using the Promega
Wizard PCR purification kit (Madison WI USA) before being
sequenced.

The H36 and H53 type strains gave two PCR bands using
primer pairs #1431/#1432 and #1417/#1418 respectively, and
15 were not sequenced.

7. Enzymes and buffers:

Restriction endonucleases and DNA T4 ligase were
purchased from Boehringer Mannheim (Castle Hill, NSW,
20 Australia). Restriction enzymes were used in the
recommended commercial buffer.

8. Sequencing of the flagellin genes:

Each PCR product was first sequenced using the
25 oligonucleotide primers used for the PCR amplification.
Primers based on the obtained sequence were then used to
sequence further, and this procedure was repeated until
the entire PCR product was sequenced.

The sequencing reactions were performed using the
30 DyeDeoxy Terminator Cycle Sequencing method (Applied
Biosystems, CA, USA), and reaction products were analysed
using fluorescent dye and an ABI377 automated sequencer
(CA, USA).

Sequence data were processed and analysed using
35 Staden programs [Sacchi CT, Zanella R C, Caugant D A,
Frasch C E, Hidalgo N T, Milagres L G, Pessoa L L, Ramos S
R, Camargo M C C and Melles C E A "Emergence of a new

- 23 -

clone of serogroup C *Neisseria meningitidis* in Sao Paulo, Brazil" J. Clin. Microbiol. 30 (1992) 1282-1286;

Staden, R.: Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing. Nucl. Acids Res. 10 (1982a) 4731-4751;

Staden, R.: An interactive graphics program for comparing and aligning nucleic acid and amino acid sequences. Nucl. Acids Res. 10 (1982b) 2951-2961;

Staden, R.: Computer methods to locate signals in nucleic acid sequences. Nucl. Acids Res. 12 (1984a) 505-519;

Staden, R.: Graphic methods to determine the function of nucleic acid sequences. A summary of ANALYSEQ options. Nucl. Acids Res. 12 (1984b) 521-538;

Staden, R.: The current status and portability of our sequence handling software. Nucl. Acids Res. 14 (1986) 217-231].

We were able to PCR amplify flagellin genes from H type strains for H7, 23, 12, 51, 45, 49, 19, 9, 30, 32, 26, 41, 15, 20, 28, 46, 31, 14, 18, 6, 34, 48, 43, 10, 52, and also from H7 strains m1004, m527, m1686, m1211, m1328, m973, m1179, m1200, m917, and m918 using primers #1575 and #1576 which are based on sequences 51-34 bp upstream and 37-54 bp downstream of start and end of the *E. coli* K-12 *fliC* gene respectively. Thus, the full sequence of the flagellin gene from these strains was obtained and the use of flanking sequence for primers makes it highly likely that they are at the *fliC* locus.

For other strains, we were only able to amplify the flagellin gene using one or more of the other three pairs of primers, which are based on sequence within the *fliC* gene, and thus only partial sequence was obtained. These amplicons may be of the *fliC* gene or one of the alternative flagellin genes. The flagellin gene sequences from H type strains for H40, 8, 21, 47, 11, 27, 35, 2, 3, 24, 37, 50, 4, 44, 38, 55, 29, 33, 5, and 56 obtained are lacking 18 and 14 codons at 5' and 3' ends respectively. The flagellin gene sequence of H39 obtained using primers

#1285/#1286 lacks 18 and 19 codons at 5' and 3' ends respectively. The flagellin gene sequence of H type strains of H17, 25 and 42 lack 23 and 21 codons at 5' and 3' ends respectively. The flagellin gene sequence of the H type strain for H54 lacks 23 and 12 codons at the 5' and 3' ends respectively. There is very little variation in the sequence at the two ends of flagellin genes and antigenic variation is due to variation in the central region of the gene. The absence of sequence for the ends of some of the flagellin genes is not important for the purpose of the present invention relating to the detection of antigenic variation by DNA sequence based means.

The *fliC* genes from H type strains of H1, H7 and H12 have been sequenced previously [Schoenhals, G. and Whitfield, C.: Comparative analysis of flagellin sequences from *Escherichia coli* strains possessing serologically distinct flagellar filaments with a shared complex surface pattern. J. Bacteriol. 175 (1993) 5395-5402] and we did not sequence the gene from the H1 strain.

We have sequenced *fliC* genes from a set of H7 strains with different O antigens, including that of *fliC* from the H7 type strain as one of the set: we have found four differences from the published H7 sequence (GenBank accession number L07388) which we believe are due to errors in the published sequence.

We have also re-sequenced the *fliC* gene from the H12 type strain, and have found one difference from the published H12 sequence (GenBank accession number L07389) which we believe is due to an error in the published sequence.

The flagellin genes from type strains H35 and H54 were also amplified using primers #1431/#1432, which are based on sequence within the *fliC* gene. Sequence data revealed that these two genes would be non-functional due to insertion sequence inserted in the middle of them. We have sequenced them to facilitate selection of primers for the functional flagellin genes.

9. Cloning of flagellin genes

DNA was digested for 2 hr at 37°C with appropriate restriction enzyme(s). The reaction product was then extracted once with phenol, and twice with ether. DNA was precipitated with 2 vols of ethanol and resuspended in water before the ligation reaction was carried out. Ligation was carried out O/N at 4°C and the ligated DNA was electroporated into one of the *E. coli fliC* mutant strains.

9.1. Cloning of flagellin genes from type strains for H1, H2, H3, H5, H6, H7, H9, H10, H11, H12, H14, H15, H18, H19, H20, H21, H24, H26, H27, H28, H29, H31, H34, H38, H39, H41, H42, H43, H45, H46, H49, H51, H52, and H56:

The full flagellin gene was PCR amplified using primers #1868 and #1870 (Table 3A). Both these primers are based on the sequences of the H7 flagellin gene of the H7 type strain. #1868 is the 5' primer: there is an *NcoI* site incorporated into the primer (Table 3B) and the flagellin gene starts at base 3 of the *NcoI* site. The 3' primer #1870 has a *BamHI* site incorporated downstream of the stop codon of the flagellin gene (Table 3B). PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, temperature varies (refer to Table 3A)/30'; extension, 72°C/1'; 30 cycles. The PCR product was purified using the Promega Wizard PCR purification kit (Madison WI USA) before being digested by restriction enzymes *NcoI* and *BamHI* and cloned into the *NcoI/BamHI* sites of plasmid pTrc99A.

Plasmid pTrc99A has a strong *trc* promoter upstream of the polylinker. Downstream of the promoter, it contains the ribosome binding site (SD sequence, see Fig 73) which is located 8bp upstream of the ATG site within the *NcoI* site. The polylinker and the SD sequence of pTrc99A are shown in Fig 73.

The plasmids generated were given pPR numbers, and

they are listed in Table 3A. In these plasmids, the expression module consists of the *trc* promoter, the SD sequence, and the full flagellin gene. The sequence of the junction region between the SD sequence and the start of flagellin gene is shown in Fig 73.

For flagellin genes from type strains for H6, H7, H9, H10, H12, H14, H18, H19, H20, H26, H28, H31, H41, H43, H45, H46, H49, H51, and H52, we have the full sequence for each gene and the primer sequences (#1868 and #1870) are conserved among them. The cloned genes therefore have the same sequence as those from the type strains.

For flagellin genes from type strains for H1, H15 and H34, we also have the full sequence. The previously published sequence of the flagellin gene from the H1 type strain was extracted from GenBank (accession number L07387) and used. Primer #1868 is conserved in all three. But, primer #1870 has the third base of the fifth last codon in the H1 sequence changed from A to G, and the third base of the second last codon changed from C to T in the H15 and H34 sequences: these changes did not change the amino acid coded, so the cloned genes encode the same gene products as those of the type strains.

For flagellin genes from type strains for H2, H3, H5, H11, H21, H24, H27, H29, H38, H39, H42, and H56, we do not have the full sequences. In the plasmids carrying genes from these type strains, the expression module consists of the *trc* promoter, the SD sequence, and the full flagellin gene with the first and the last 21 base pairs being determined by the primer sequences which are based on the H7 flagellin gene of the H7 type strain. The sequence of the junction region between the SD sequence and the start of flagellin gene is shown in Fig 73.

9.2. Cloning of the flagellin gene from type strain of H23:

The full flagellin gene was PCR amplified using primers #1868 and #1869 (Table 3A). #1868 is the 5'

primer: there is an *Nco*I site incorporated into the primer (Table 3B) and the flagellin gene starts at base 3 of the *Nco*I site. The 3' primer #1869 has a *Sal*I site incorporated downstream of the stop codon of the flagellin gene (Table 3B). PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, 55°C/30'; extension, 72°C/1'; 30 cycles. The PCR product was purified using the Promega Wizard PCR purification kit (Madison WI USA) before being digested by restriction enzymes *Nco*I and *Sal*I and cloned into the *Nco*I/*Sal*I sites of plasmid pTrc99A to give plasmid pPR1942.

Plasmid pTrc99A has a strong *trc* promoter upstream of the polylinker. Downstream of the promoter, it contains the ribosome binding site (SD sequence, see Fig 73) which is located 8bp upstream of the ATG site within the *Nco*I site. The polylinker and the SD sequence of pTrc99A are shown in Fig 73.

The expression module of pPR1942 consists of the *trc* promoter, the SD sequence, and the full flagellin gene. The sequence of the junction region between the SD sequence and the start of flagellin gene is shown in Fig 73.

9.3. Cloning of flagellin genes from type strains of H30, H32 and H33:

The full flagellin gene was PCR amplified using primers #1868 and #1871 (Table 3A). #1868 is the 5' primer: there is an *Nco*I site incorporated into the primer (Table 3B) and the flagellin gene starts at base 3 of the *Nco*I site. The 3' primer #1871 has a *Pst*I site incorporated downstream of the stop codon of the flagellin gene (Table 3B). PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, temperature varies (refer to Table 3A)/30'; extension, 72°C/1'; 30 cycles. The PCR product was purified using the Promega Wizard PCR purification kit (Madison WI USA) before being digested by restriction enzymes *Nco*I and *Pst*I

and cloned into the *NcoI*/*PstI* sites of plasmid pTrc99A.

Plasmid pTrc99A has a strong *trc* promoter upstream of the polylinker. Downstream of the promoter, it contains the ribosome binding site (SD sequence, see Fig 73) which is located 8bp upstream of the ATG site within the *NcoI* site. The polylinker and the SD sequence of pTrc99A are shown in Fig 73.

For flagellin genes from type strains for H30 and H32, we have the full sequence. Primer #1868 sequence is conserved in both of them. But, primer #1871 has the third base of the fourth last codon in both sequences changed from G to A to remove a *PstI* site (see Table 3B): this change did not change the amino acid coded. The expression module consists of the *trc* promoter, the SD sequence, and the full flagellin gene coding for a gene product which is same as that of the type strain. The sequence of the junction region between the SD sequence and the start of flagellin gene is shown in Fig 73.

We do not have the full sequence for the flagellin gene from the H33 type strain. In the plasmid containing the H33 type strain gene, the expression module consists of the *trc* promoter, the SD sequence, and the full flagellin gene with the first and the last 21 base pairs been determined by the primer sequences which were used for the cloning of H30 and H32. The sequence of the junction region between the SD and the start of flagellin gene is shown in Fig 73.

9.4. Flagellin genes from type strains for H4 and H17:

For the flagellin genes of H4 and H17 type strains the full sequence was not obtained, and the sequenced parts were PCR amplified and cloned into plasmid pPR1951 to give in each case a gene in which the first 26 and the last 31 codons are based on the sequence of the H7 flagellin gene of the H7 type strain.

9.4.1

Construction of expression plasmid vector

pPR1951:

The first 26 codons of the H7 flagellin gene was first PCR amplified using primers #1868 and #1872 (Table 3B). #1868 is the 5' primer: there is an *NcoI* site incorporated into the primer (Table 3B) and the flagellin gene starts at base 3 of the *NcoI* site. Primer #1872 was made to have the last two codons (codons 25 and 26) changed from CTG TCG (Leucine and Serine) to GGA TCC (Glycine and Serine) to generate a *BamHI* site. This PCR fragment was digested with *NcoI* and *BamHI* before being cloned into the *NcoI/BamHI* sites of pTrc99A to make plasmid pPR1949.

The last 31 codons (including the stop codon) of the H7 flagellin gene was PCR amplified using primers #1884 and #1871 (Table 3A). The 5' primer, #1884, has the first two of the 31 codons changed from TCG AAA (Serine and Lysine) to TCT AGA (Serine and Arginine) to generate a *XbaI* site (Table 3B). The 3' primer #1871 has a *PstI* site incorporated downstream of the stop codon (Table 3B). This PCR fragment was digested with *XbaI* and *PstI*, and then cloned into the *XbaI/PstI* sites of pPR1949 to make plasmid pPR1951.

9.4.2 Cloning of flagellin genes from the H4 and H17 type strains:

The central regions of flagellin genes from type strains H4 and H17 were PCR amplified using primers #1878 and #1885 (Table 3B), which have a *BamHI* and a *XbaI* incorporated at their ends respectively. PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, 65°C/30'; extension, 72°C/1'; 30 cycles. The PCR product was purified using the Promega Wizard PCR purification kit (Madison WI USA) before being digested by restriction enzymes *BamHI* and *XbaI* and cloned into the *XbaI/BamHI* sites of plasmid pPR1951 to make plasmids pPR1955 (H4) and pPR1957 (H17).

The expression module of plasmids pPR1955 and pPR1957

consists of the *trc* promoter, the SD sequence, the first 24 codons of the H7 flagellin gene (of the H7 type strain), 2 codons encoding Glycine and Serine, 292 or 293 codons of the central region based on the flagellin gene obtained from the H4 or H17 type strain respectively, 2 codons encoding Serine and Arginine, and then the last 29 codons of the H7 flagellin gene (of the H7 type strain).

10. *Expression of flagellin gene plasmids in E. coli strains lacking the fliC gene, and identification of the H antigens encoded by these plasmids:*

Plasmids carrying flagellin genes as described in section 9 (see Table 3A for a list) were electroporated into strains M2126 or P5560. Strains M2126 and P5560 do not have functional *fliC* genes, and are not motile when examined under a microscope. Transformants carrying any of the plasmids listed in Table 3A are motile when examined under a microscope. Thus, the flagellin genes in all of the plasmids are expressed.

The antigenic specificity of the flagellin of each transformant was then determined by slide agglutination.

10.1 *Flagellin genes from type strains for H2, H5, H6, H7, H9, H11, H14, H15, H18, H19, H20, H21, H23, H24, H26, H27, H28, H29, H30, H31, H32, H33, H34, H39, H41, H42, H43, H45, H46, H49, H51, H52, and H56:*

As shown in Table 3A, strains with plasmids carrying these flagellin genes expressed the same H antigen as their respective flagellin parent strain.

For flagellin specificities H2, H5, H6, H7, H9, H14, H15, H18, H19, H20, H23, H24, H26, H27, H28, H29, H31, H33, H39, H51, H52, and H56, there was no cross reaction reported between these flagellins and flagellin antisera for other H antigens [Ewing, W. H.: *Enterobacteriaceae*, Elsevier Science Publishers, Amsterdam, The Netherlands, 1986], and we conclude that we have in each case sequenced the gene

- 31 -

encoding the flagellin of the expected specificity from the respective type strain.

It has been observed that cross reactions exist between some type strains and certain antisera at different levels of dilution (of the antisera), being H11 with anti-H21 and anti-H40, H21 with anti-H11, H30 with anti-H32, H32 with anti-H30, H34 with anti-H24 and anti-H31, H41 with anti-H37 and anti-H39, H42 with anti-H6, H43 with anti-H37, H45 with anti-H20, H46 with anti-H17, and H49 with anti-H39 [Ewing, W. H.: Edwards and Ewing's identification of the *Enterobacteriaceae.*, Elsevier Science Publishers, Amsterdam, The Netherlands, 1986]. We have tested strain M2126 or strain P5560 carrying plasmids containing flagellin genes obtained from each of these type strains (H11, H21, H30, H32, H34, H41, H42, H43, H45, H46, and H49) with the appropriate cross-reacting antisera.

For strain M2126 or strain P5560 carrying plasmids containing flagellin genes obtained from type strains H11, H34, H41, H42, H43, H45, H46, and H49, no cross reaction was found. We conclude that we have in each case sequenced the gene coding the flagellin of the expected specificity from the respective type strain.

Cross reaction was observed for strain P5560 carrying plasmid pPR1948 (containing the flagellin gene obtained from the H30 type strain) with anti-H32 serum, strain P5560 carrying pPR1940 (containing the flagellin gene obtained from the H32 type strain) with anti-H30 serum, and strain M2126 carrying plasmid pPR1995 (containing the flagellin gene obtained from the H21 type strain) with anti-H11 serum.

We note that the reported cross reactions between the H30 type strain and anti-H32, the H32 type strain and anti-H30, and the H21 type strain and anti-H11 happened at a higher level of dilution (of antisera) than for all other type strains with the antisera mentioned above [Ewing, W. H.: Edwards and Ewing's identification of the

Enterobacteriaceae., Elsevier Science Publishers, Amsterdam, The Netherlands, 1986]. We conclude that except for these three cases, the antiserum used were supplied at a dilution which did not exhibit cross reactions. For the three strains carrying flagellin genes cloned from type strains for H30, H32, and H21, it was necessary to further dilute the antiserum.

Strain P5560 carrying plasmid pPR1948 (containing the flagellin gene obtained from the H30 type strain) agglutinates with anti-H30 serum when the antiserum is diluted to 1:60, but agglutinates with anti-H32 serum only at a dilution of 1:10 and not at a 1:20 dilution (note that the antisera we used have been diluted before reaching our hands). In contrast, strain P5560 carrying plasmid pPR1940 (containing flagellin gene obtained from the H32 type strain) agglutinates with anti-H32 serum when the antiserum is diluted at 1:100, but agglutinates with anti-H30 serum only at a 1:10 dilution and not at a 1:10 dilution. Thus, we conclude that the flagellin genes we sequenced from type strains for H30 and H32 encode flagellins of H30 and H32 specificities respectively.

Strain M2126 carrying plasmid pPR1995 (containing the flagellin gene obtained from the H21 type strain) agglutinates with anti-H21 serum when the antiserum is diluted to 1:40, but agglutinates only with undiluted anti-H11 serum and not at a 1:10 dilution (note that the antisera we used have been diluted before reaching our hands). In contrast, strain M2126 carrying plasmid pPR1981 (containing flagellin gene obtained from the H11 type strain) did not agglutinate with anti-H21 serum. Thus, we conclude that the flagellin genes we sequenced from type strains for H21 encodes flagellin of H21 specificity.

10.2 *Flagellin genes from type strains of H1 and H12:*

These two genes are very similar in sequence, with 8 a.a difference between the gene products. It has been

known that some cross-reaction exists between anti-H12 serum and the H1 type strain and between anti-H1 serum and the H12 type strain [Ewing, W. H.: Edwards and Ewing's identification of the *Enterobacteriaceae*., Elsevier Science Publishers, Amsterdam, The Netherlands, 1986]. Strain M2126 carrying pPR1920 (carrying a flagellin gene from the H1 type strain, Table 3A) agglutinates with anti-H1 serum when the antiserum is diluted to 1:100, but agglutinates only with undiluted anti-H12 serum and not at a 1:10 dilution (please note that the antisera we used have been diluted before reaching our hands). In contrast, strain M2126 carrying plasmid pPR1990 (carrying a flagellin gene from the H12 type strain, Table 3A) agglutinates with anti-H12 serum when the antiserum is diluted at 1:100, but agglutinates only with undiluted anti-H1 serum and not at a 1:10 dilution. Thus, we conclude that the flagellin genes we sequenced from type strains for H1 and H12 encode flagellins of H1 and H12 specificities respectively.

10.3. Flagellin gene coding for H16:

Strain P5560 carrying plasmid pPR1969 agglutinated with anti-H16 serum. pPR1969 carries a flagellin gene amplified from the H3 type strain. It has been shown that this H3 type strain is a biphasic strain which can express H3 and H16 specificities [Ratiner, Y. A. (1985) "Two genetic arrangements determining flagellar antigen specificities in two diphasic *E. coli* strains. FEMS Microbiol Lett 19: 317-323]. Thus, the H3 type strain has two flagellin genes coding for H3 and H16 specificities. We conclude that we have cloned and sequenced the H16 flagellin gene from this H3 type strain.

10.4 Flagellin gene coding for H4:

The flagellin genes obtained from type strains for H4 and H17 are nearly identical, with 4 a.a. difference in the gene products. Plasmid pPR1955 carries a flagellin

gene from the H4 type strain, and plasmid pPR1957 carries a flagellin gene from the H17 type strain. Strain P5560 carrying plasmid pPR1955 or plasmid pPR1957 agglutinated with anti-H4 serum, but not with anti-H17 serum. It has been shown that the type strain for H17 is a biphasic strain which can express H17 and H4 [Ratiner, Y. A. (1985) "Two genetic arrangements determining flagellar antigen specificities in two diphasic *E. coli* strains. FEMS Microbiol Lett 19: 317-323]. The flagellin gene obtained from type strain for H44 is also highly similar to that obtained from the H4 type strain, with 2 a.a. difference in the gene products. It has been shown that the H44 type strain has two complete flagellin genes, being H4 and H44 [Ratiner, Y. A. (1998) "New flagellin specifying genes in some *E. coli* strains" J. Bacteriol 180: 979-984]. Thus, we conclude that all the three flagellin genes (obtained from type strains for H4, H17 and H44, and sequenced) encode the H4 flagellin, and that the flagellin genes for H17 and H44 specificities have not yet been cloned.

10.5 *Flagellin gene coding for H10:*

The flagellin genes obtained from type strains for H10 and H50 are nearly identical, with 3 a.a. difference in the gene products. Strain P5560 carrying plasmid pPR1923 (which carries a flagellin gene from the H10 type strain) agglutinated with anti-H10 serum. We conclude that the sequence obtained from the H10 type strain encodes the H10 flagellin. It is not clear if the sequence obtained from the H50 type strain encodes H10 or H50 (see below section for H50).

10.6 *Flagellin gene coding for H38:*

The flagellin genes obtained from type strains for H38 and H55 are nearly identical, with only 1 a.a. difference in the gene products. Strain M2126 carrying plasmid pPR1984 (carrying the flagellin gene from the type strain H38) agglutinated with anti-H38 serum, but not with

anti-H55 serum. It also has been shown that the type strain for H55 has two complete flagellin genes coding for H55 and H38 specificities [Ratiner, Y. A. (1998) "New flagellin specifying genes in some *E. coli* strains" J. Bacteriol 180: 979-984]. Thus, we conclude that both cloned genes encode the H38 flagellin.

10.7 Summary:

Flagellin genes coding for 39 H antigens have been identified, being those for specificities H1, H2, H4, H5, H6, H7, H9, H10, H11, H12, H14, H15, H16, H18, H19, H20, H21, H23, H24, H26, H27, H28, H29, H30, H31, H32, H33, H34, H38, H39, H41, H42, H43, H45, H46, H49, H51, H52, and H56.

11. Comparison and alignment of the flagellin genes:

Programs Pileup [Devereux, J., Haeberli, P. and Smithies, O.: A comprehensive set of sequence analysis programs for the VAX. Nucl. Acids Res. 12 (1984) 387-395] and Multicomp [Reeves, P.R., Farnell, L. and Lan, R.: MULTICOMP: a program for preparing sequence data for phylogenetic analysis. CABIOS 10 (1994) 281-284] were used.

The previously published sequence of H1 (GenBank accession number L07387) was extracted from GenBank and used. Because we did not sequence H36 and H53 flagellin genes and we did not have the H16 type strain, we only compared 51 flagellin genes of H type strains and the *fliC* genes from the additional 10 H7 strains.

Among the H7 *fliC* genes, the percentage of DNA difference ranged from 0.0 to 2.39%. The flagellin genes from type strains for H40 and H8 are identical. Some others are nearly identical: H21 and H47 (1.5% difference), H12 and H1 (2.6% difference), H10 and H50 (0.3% difference), H38 and H55 (0.1% difference), H4, H44 and H17 are very similar, the pairwise difference ranging from 0.33% to 0.87%.

For the flagellin genes obtained from type strains for H4, H17 and H44, we have shown that all the three genes encode flagellin with the H4 specificity (see above). For the flagellin genes obtained from type strains fro H21 and H47, and H38 and H55, we have confirmed the specificities for one for each pair and have good reason to conclude that both genes of each pair encode the same H specificity (see above section), being that for H21 and H38 specificities respectively.

For the flagellin genes obtained from type strains for H10 and H50, we have confirmed that the one from the H10 type strain encodes H10 specificity. As these two genes are highly similar, we have presumed that they both encode H10 specificity.

In the cases where the flagellin gene from two type strains is near identical, we conclude that both genes code for flagellin of the same H specificity and that one or other strain has an additional locus which carries the functional gene, although the flagellin genes sequenced do not appear to be mutated.

We have shown by cloning and expression that the flagellin genes obtained from the H1 and H12 type strains encode H1 and H12 specificities respectively (see above section). The neucleotide difference between these two genes is higher at 2.6% (see above), but still within the normal range for variation within a gene in *E. coli*. The two antigens cross react, and this cross reaction must be due to the high level similarity of the flagellins encoded by these two genes.

As discussed above, genes encoding some H antigens have been shown to be located at loci other than *fliC*. H3, H36, H47, H53 have been shown to be at a locus called *flkA*, H44 and H55 at *fllA*, and H54 at *flmA* [Ratiner Y A (1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984]. However, these strains may carry a *fliC* in addition to *flkA*, *fllA* or *flmA* [Ratiner Y A (1998) "New flagellin-specifying genes in some

Escherichia coli strains" J. Bacteriol. 180 979-984].

The flagellin gene encoding H48 was previously sequenced from *E. coli* strain K-12 [Ku wajima G, Asaka J, Fujiwara T, Node K and Kondo E "Nucleotide sequence of the hag gene encoding flagellin of *Escherichia coli*" J Bacteriol. 168 (1986) 1479-1483]. We have sequenced the *fliC* gene from the H48 type strain, and found that it is identical to that from K-12.

The H54 gene is known to be at *flmA* [Ratiner Y A (1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984] and the finding of a non-functional presumptive *fliC* locus in the H54 strain shows that it is present but not expressed. However, we have not amplified and sequenced the functional *flmA* gene of this strain.

Using the 43 unique sequences (being the 39 identified genes with confirmed specificities and the flagellin genes obtained from the H8 (or H40), H25, H37, and H48 type strains) and the sequences from the two non-functional flagellin genes (from H type strains H35 and H54) (see Table 3) we have been able to determine antigen specific primers for each of the H antigen specificities and thereby show that it is practicable to detect *E.coli* strains carrying specific H antigens without false positives from strains of other H types. There is no reason to expect that the addition of 11 sequences to the 43 unique sequences obtained will affect the general conclusion, as unlike previous reports, our study covers flagellin sequences for a substantial majority of known *E. coli* H antigen specificities.

Our study of 11 H7 genes from strains of eight different O antigens shows limited variation which was such that the variation within genes for H antigens does not affect the ability to select antigen specific primers. O:H combinations in general define a strain and as some of the strains thus defined were quite distant from each other in a study by Whittam [Whittam T S, wolfe M L,

- 38 -

Wachsmuth I K, Orskov I and Wilson R A "Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea" Infect. Immun. 61 (1993) 1619-1629] the variation we observe is thought to represent that generally present in H7 genes. We also obtained more than one sequences for flagellin genes for H specificities H4, H10, and H38, and again the level of variation within a given specificities is very low. However, there is a low possibility that primers chosen without knowledge of the variation within genes of each H specificity could fail to give positive results with some isolates due to chance choice of primers which cover a base or bases which contribute to this low level variation. The variation within the H7 genes is in the normal range for variation within a gene in *E. coli* and if this possibility did occur it would be easy to use an alternate primer pair. For example, if a first primer in a primer pair is unable to hybridise to a target region because of low level variation in that region, a positive result may be achieved by using a second primer in that pair together with a third primer, whether or not the third primer is specific for the flagellin gene. Where the third primer is not specific for the flagellin gene, the specificity of the primer pair derives from the specificity of the second primer. The observation that the overall level of variation within gene for a given H specificity is very low making it extremely unlikely that the regions covered by the two primers specific for H specificity would both have undergone change in the same strain.

There are 54 known H antigens for *E. coli* and of these there are 11 H antigen specificities for which we do not as yet have sequence. It will be easy to determine these sequences and determine primer pairs specific for these H antigens by comparing these sequences with the 45 obtained sequences (see Table 3), and also modify the primers selected for any H antigen for which we already

- 39 -

know the sequence in the unlikely event that there is a possibility of false positives with the primers selected.

The sequences for the remaining H antigens can be obtained in one of the following ways:

5

10

15

20

25

1. where we have two bands by PCR (H36 and H53 type strains), we purify each and sequence, and also clone each into a strain mutated in its *fliC* gene and determine the H antigen expressed by use of specific sera. In this way a specific sequence can be related to an H antigen specificity. The other band which represents an H antigen gene for a different specificity is expected to include a mutant gene or a gene similar to one of those for a known H specificity, but if not may represent a new specificity for which primer pairs could be selected. It may be difficult to obtain expression of flagellin genes when cloned from *E. coli* due to cloning together with regulatory sequences which prevent expression. This is easily avoided by cloning the major segment of the gene into a functioning *fliC* gene to replace the equivalent segment of that gene, using standard site directed mutagenesis to give suitable restriction sites within the cloned gene and incorporating those restriction sites into primers used to amplify the major segment of the gene to be studied to facilitate the cloning. We have cloned and sequenced the PCR bands from the H36 and the H55 type strains using this method (see section 16).

30

35

2. Where two or more strains have the same flagellin gene sequence, the genes are cloned as above and the H antigen specificity represented by this sequence is determined. This identifies the strain in which the expected gene is expressed and also those strains for which we have sequenced a gene which is not being expressed. We then clone the gene for the antigen expressed in these strains by making a bank of plasmid clones using chromosomal DNA and select for a clone which

- 40 -

is expressing an H antigen different from the one represented by the known sequence. This can be done by taking advantage of the fact that the H antigen is on flagellin, the protein of the bacterial flagellum used for movement of the bacteria. In the presence of antibodies specific to that flagellum the bacteria cannot swim. For selection the clones are placed in a situation in which motile cells can swim away from the others and be collected. There are many versions of these techniques and any could be used. One version is to place the bacteria on a nutrient agar plate with reduced agar content such that bacteria can swim away from the site of inoculation. This is easily seen as growth on the plate and a sample of the bacteria which are motile can be recovered and cultivated. In this way bacteria carrying cloned H antigen genes can be selected. If the medium in the plate has antibody added to it only bacteria which express an H antigen different to that recognised by the antiserum will be able to swim. Specifically if the antiserum used is specific for the H antigen expressed by the gene for which we have sequence, only clones which express a different H antigen, such as those expressing the H antigen expressed by the H type strains used to make the plasmid, will be selected. Once the clone is obtained, the H antigen gene can be sequenced.

Our work has shown that there are at least 7 cases where the H antigen type strains carry two H antigen genes which appear to be complete and have the potential to function. However, while *E. coli* does not (in general) have a capacity to express more than one flagellin gene, it is striking that there are several loci for flagellin genes [Ratiner Y A (1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984]. Several of the pairs of H type strains with identical or near identical sequence do not include any of the H antigen types shown by Ratiner [Ratiner Y A

(1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984] to map other than at *fliC* although these predominate. This suggests that there are additional cases where the expressed gene is not the only flagellin gene present. However the fact that many of the cases where we obtained flagellin genes of identical or near identical sequence and/or two flagellin genes from one strain involve type strains found by Ratiner [Ratiner Y A (1998) "New flagellin-specifying genes in some *Escherichia coli* strains" J. Bacteriol. 180 979-984] to map away from *fliC* are among those near identical to others, indicates that the phenomenon is of limited extent. Nonetheless it remains possible even where only one gene has been obtained by PCR, that it is one of a pair of flagellin genes, the other not being amplified by the primers used, and further that it is the one not amplified which is expressing the H antigen of the strain.

It will therefore be necessary to clone as described above each of the flagellin genes we have sequenced and confirm that it expresses the expected antigen to ensure that the invention give results corresponding to those of the traditional serotyping scheme. In the event that it does not, the gene for the type antigen can be cloned and sequenced by the means described above.

The 11 H7 *fliC* sequences fell into three groups, one comprising the genes from the O157:H7 and O55:H7 strains, which were identical, as expected given the proposed relationship between the clones. It has been shown that *E. coli* O157:H7 and O55:H7 clones are closely related [Whittam T S, wolfe M L, Wachsmuth I K, Orskov I and Wilson R A "Clonal relationships among *Escherichia coli* strains that cause hemorrhagic colitis and infantile diarrhea" Infect. Immun. 61 (1993) 1619-1629] thus it was expected that the H7 *fliC* genes from O157 and O55 would be identical. Among the H7 *fliC* sequences, we can identify primers specific to the H7 *fliC* gene for each of the three H7 groups. Two of these primers in combination with an H7

specific primer gave two primer pairs specific for the H7 gene of from the O157:H7 and O55:H7 clones.

13. *Specific oligonucleotide primers for each of the 43 flagellin genes*

Two oligonucleotide primers were chosen based on each of the 43 sequences. None of them had more than 85% identity with any other of 61 flagellin gene sequences. Thus, these primers are specific for each H type. These primers are listed in Table 3.

The flagellin gene of the H54 type strain is a mutated gene. It has an insertion sequence (IS1222) inserted into a normal flagellin gene of H21. Thus, primers for H21 would amplify a fragment of different size in H54. We also provide 2 primers based on the insertion sequence (see H54 row in Table 3), and the use of one of them in combination with one of the H21 primers will generate a PCR band only in H54, which will also differentiate those strain carrying the mutated H21 gene from those expressing the H21 flagellin gene.

The *flic* gene of H35 type strain is also a mutated gene. It has an insertion sequence (IS1) inserted into a normal flagellin gene of H11. Thus, primers for H11 would amplify a fragment of different size in H35. We also provide 2 primers based on the insertion sequence (see H35 row in Table 3), and the use of one of them in combination with one of the H11 primers will generate a PCR band only in H35, which will also differentiate those strain carrying the mutated H11 gene from those expressing the H11 flagellin gene.

14. *Testing of the H7 specific oligonucleotide primers*

Primer pair #1806/#1809 (see Table 3) was used to carry out PCR on chromosomal DNA samples of all the 54 H type strains and the H7 strains listed in Table 1. PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, 58°C/30'; extension,

72°C/1'; 30 cycles. PCR reaction was carried out in an volume of 50ul for each of the chromosomal sample. After the PCR reaction, 5µl PCR product from each sample was run on an agarose gel to check for amplified DNA.

5 Primer pairs #1806/#1809 produced a band of predicted size with all the 11 strains expressing H7, but gave no band with other H type strains. Thus, these primers are H7 specific.

10 15. Testing of oligonucleotide primers specific to H7 of O157 and O55:

15 Based on a comparison of the *fliC* sequences of 11 different H7 strains, we have identified two oligonucleotides [#1696 (5'-GGCCTGACTCAGGCGGCC) at positions 178 to 195 in M527 and #1697 (5'-GAGTTACCGGCCTGCTGA) positions 1700-1683 in M527] which are unique to H7 of O157 and O55. Although not identical to any parts of the *fliC* sequences of any other H7 strains, these two primers are identical or have high level
20 similarity to *fliC* genes of some other H types. However a combination of one of these primers with one of the H7 specific primers can give specificity for H7 of O157:H7 and O55:H7 *E. coli*.

25 Primer pairs #1696/#1809 and #1697/#1806 were used to carry out PCR on chromosomal DNA samples of all the H type strains and the H7 strains listed in Table 1. PCR reactions were carried out under the following conditions: denaturing, 94°C/30'; annealing, 61°C/30' (for #1696/#1809) or 60°C/30' (for #1697/#1806); extension,
30 72°C/1'; 30 cycles. PCR reaction was carried out in an volume of 50µl for each of the chromosomal samples. After the PCR reaction, 5µl PCR product from each sample was run on an agarose gel to check for amplified DNA.

35 Both primer pairs produced a band of predicted size with both of the O157:H7 strains (strains M1004 and M527, see Table 1), and the O55:H7 strain (strain M1686, see Table 1), but gave no band with other strains. Thus, these

- 44 -

two pairs of primers are specific to H7 genes of O157:H7 and O55:H7 *E. coli* strains.

5 16. Identification of flagellin genes for the remaining 15 *H* specificities.

16.1. Sequencing the potential *flkA* gene coding for the H36 flagellin:

10 Using primers #1431 (5'- atg gca caa gtc att aat acc caa c) and #1432 (5'- cta acc ctg cag cag aga ca), we have amplified two bands from the H36 type strain. PCR reaction was carried out under the following conditions: denaturing, 94oC/30'; annealing, 57oC/30'; extension, 72oC/1'; 30 cycles. These two PCR fragments were then
15 cloned into the pGEM-T vector using the Promega pGEM-T cloning kit (Madison WI USA) to make plasmids pPR1992 and pPR1993. Inserts from both plasmids were first sequenced using the M13 universal primers (which bind to the pGEM-T DNA flanking the insertion site). For pPR1992, primers
20 based on the sequence obtained were then used to sequence further, and this procedure was repeated until the insert was fully sequenced.

The sequence of the insert of pPR1992 is identical to that of the H12 flagellin gene sequence except perhaps for the first 8 and last 7 codons which are encoded by the PCR
25 primers in plasmid pPR1992. We have only sequenced the two ends of the insert of plasmid pPR1993 (Figures 71 and 72), and the sequences of the two ends of the insert of pPR1993 are very similar to ends of other sequenced
30 flagellin genes. We conclude that the insert of plasmid pPR1993 encodes a flagellin gene. The full sequence of the insert of plasmid pPR1993 can be obtained using the same method as for the sequencing of the insert of plasmid pPR1992. It is known that *flkA* gene encodes the H36
35 flagellin [Ratiner, Y. A. (1998) "New flagellin specifying genes in some *E. coli* strains" J. Bacteriol 180: 979-984], and it is highly likely that plasmid pPR1993 contains the

flkA gene of the H36 type strain. H specificities can be confirmed by slide agglutination.

The currently uncharacterised sequence of both ends and of DNA flanking these two sequenced genes can be obtained by PCR walking and sequencing. Methods for PCR walking from a known sequence to an unknown region in chromosomal DNA are available (see [Siebert, P. D. , A. Chenchi, D. E. Kellogg, A. Lukyanov and S. A. Lukyanov (1995) "An improved PCR method for walking in uncloned genomic DNA." Nuc. Acids Res. 23: 1087-1088])).

The sequenced genes then can be PCR amplified and cloned using the method(s) described in section 9. Flagellins expressed by strain M2126 carrying these plasmids then can be determined by use of specific sera.

The sequences flanking the *flkA* gene can then be used to PCR amplify other *flkA* genes (see below).

16.2 The *flkA* genes coding for H3, H47 and H53:

It has been shown that flagellins H3, H47 and H53 are encoded by *flkA* genes in the type strains [Ratiner, Y. A. (1998) "New flagellin specifying genes in some *E. coli* strains" J. Bacteriol 180: 979-984]. These genes can be PCR amplified using primers based on the sequences flanking the *flkA* gene in the H36 type strain. These PCR fragments can then be sequenced, and the genes expressed in strain M2126 for the identification of these genes.

16.3 The *fllA* genes coding for H44 and H55:

It is known that flagellins H44 and H55 are coded by *fllA* genes.

16.3.1 The H55 flagellin gene:

Using primers #1868 and #1870 (Table 3B), we have amplified two bands from the H55 type strain. PCR reaction was carried out under the following conditions: denaturing, 94°C/30'; annealing, 50°C/30'; extension, 72°C/1'; 30 cycles. These two PCR fragments were then

cloned into the pGEM-T vector using the Promega pGEM-T cloning kit (Madison WI USA) to make plasmids pPR1994 and pPR1989. Inserts from both plasmids were first sequenced using the M13 universal primers (which bind to the pGEM-T DNA flanking the insertion site). Primers based on the sequence obtained were then used to sequence further, and this procedure was repeated until both inserts were fully or partly sequenced.

The sequence of the insert of pPR1994 is highly similar to that of the flagellin gene of the H38 type strain, with 1 amino acid difference in the gene products. We have only sequenced the two ends of the insert of plasmid pPR1989 (figures 70A and 70B), and the sequences of the two ends of the insert of pPR1989 are very similar to ends of other sequenced flagellin genes. We conclude that the insert of plasmid pPR1989 encodes a flagellin gene. The full sequence of the insert of plasmid pPR1989 can be obtained using the same method as for the sequencing of the insert of plasmid pPR1994. It is known that the H55 type strain carries flagellin genes for both H38 and H55, and that the H55 flagellin gene is at the *flhA* locus [Ratiner, Y. A. (1998) "New flagellin specifying genes in some *E. coli* strains" J. Bacteriol 180: 979-984]. Thus, it is highly likely that plasmid pPR1989 contains the *flhA* gene of the H55 type strain.

The currently uncharacterised sequence of both ends and of DNA flanking these two sequenced genes can be obtained by PCR walking and sequencing. Methods for PCR walking from a known sequence to an unknown region in chromosomal DNA are available (see [Siebert, P. D. , A. Chenchi, D. E. Kellogg, A. Lukyanov and S. A. Lukyanov (1995) "An improved PCR method for walking in uncloned genomic DNA." Nuc. Acids Res. 23: 1087-1088]).

The sequenced genes then can be PCR amplified and cloned using the method(s) described in section 9. Flagellins expressed by strain M2126 carrying these plasmids then can be determined by use of specific sera.

16.3.2 *The H44 flagellin gene:*

The sequence information for DNA flanking the *flhA* gene in the H55 type strain can then be used to PCR, sequence and identify the *flhA* gene in the H44 type strain.

16.4 *The flmA gene coding for H54:*

This gene can be cloned by making a bank of plasmid clones in strain M2126 using chromosomal DNA of the H54 type strain and selecting for a transformant which is motile on an agar plate. This is done by taking advantage of the fact that the H antigen is on flagellin, the protein of the bacterial flagellum used for movement of the bacteria. Strain M2126 lacks flagellin. Once the clone(s) is obtained and identified by use of anti-H54 serum, the flagellin gene can be sequenced. It is possible that clones expressing different flagellin specificities can be obtained, and each of them can be identified by using different sera.

16.5 *The flagellin genes obtained from the H37 and H48 type strains:*

We have used primers #1868 and #1869 (both were based on the sequence obtained from the H48 type strain, also see section 9) and primers #1868 and #1870 (both were based on the sequences of the H7 flagellin gene of the H7 type strain, also see section 9) to PCR amplify and clone the sequenced flagellin genes from the H48 and H37 type strains respectively. Strain P5560 carrying the plasmid containing either the cloned gene was not motile and did not react with the appropriate antisera. It is highly likely that mutations have occurred due to PCR errors. This can be resolved by re-amplification and re-cloning of the genes.

16.6 *The flagellin gene obtained from the H25 type*

strain:

The flagellin gene sequence we first obtained from the H25 type strain lacks 23 and 21 codons at 5' and 3' ends respectively. We could not amplify the full gene from the H25 type strain using primers based on the H7 flagellin gene of the H7 type strain, and it was necessary to get the full sequence of this flagellin gene by other means.

We have used primers (#2650: 5' - cag cga tga aat act tgc cat and #2648: 5' - caa tgc ttc gtg acg cac) based on the genes (*fliD* and *fliA* respectively) flanking *fliC* gene in *E. coli* K-12 [Blattner, F. R., G. I. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley and et al. (1997) "The complete genome sequence of *E. Coli* K12" *Science* 277: 1453-1474] and primers (#2658: 5' - gcc tga gtc aga cct ttg and # 2653 5' - aac ctg tct gaa gcg cag) based on the flagellin sequence obtained from the H25 type strain to PCR amplify both ends of the flagellin gene. The PCR product was then sequenced, and we have now obtained the full flagellin gene sequence and sequence for the DNA flanking the flagellin gene from type strain H25 (Figure 69). Now, it is straightforward to PCR amplify, clone and express, and identify this gene using the methods described in sections 9 and 10.

16.7 *The flagellin genes obtained from the H8 and H40 type strains:*

The flagellin gene sequences obtained from both the H8 and H40 type strains lack 18 and 15 codons at 5' and 3' ends respectively. We have used primers based on the H7 flagellin gene of the H7 type strain to PCR amplify and clone the full genes from these two strains. Strain M2126 carrying plasmid made this way was not motile under microscope and did not react with the appropriate antisera. This could be due to PCR errors as mentioned in section 16.5 or perhaps the first and last few amino acids encoded by the primers (based on H7 flagellin gene) are

uncompatible in this case.

The full sequence of the full gene can be obtained using method described in section 16.6. The flagellin gene can then be PCR amplified, cloned and expressed, and identified using the methods described in sections 9 and 10.

The gene products of the flagellin genes obtained from the H8 and H40 type strains are identical. Thus, one of these two H specificities must be encoded by a unknown gene, and it can be cloned and identified using the method described in the section 16.8.

16.8 *Flagellin genes coding for H17, H35, and H50:*

As mentioned above, the sequenced flagellin genes from the H17 and H50 type strains encode H4 and H10 specificities respectively. The flagellin gene sequence obtained from the H35 strain has a insertion and encodes a non-functional gene (see section 8). Thus, genes coding for these flagellins have not been identified, and their location is unknown. One can use primers based on DNA flanking *fliC*, *fllA*, *flkA*, and *flmA* to do PCR on the type strain for each of the flagellin antigen. PCR products can then be sequenced, and possible genes can be cloned, expressed and identified then.

If the target gene is not PCR amplified using primers based on sequence of these loci or sequence flanking these loci, it can be cloned by making a bank of plasmid clones in strain M2126 using chromosomal DNA of the type strain and selecting for a transformant which is motile on an agar plate. This is done by taking advantage of the fact that the H antigen is on flagellin, the protein of the bacterial flagellum used for movement of the bacteria. Strain M2126 lacks flagellin. Once the clone(s) is obtained and identified by use of antisera, the flagellin gene can be sequenced. It is possible that clones expressing different flagellin antigens can be obtained,

and each of them can be identified by using different antisera. Antiserum for H50 can be prepared using standard methods [Ewing, W.H.: Edwards and Ewing's identification of the *Enterobacteriaceae*., Elsevier Science Publishers, Amsterdam, The Netherlands, 1986].

O antigen

Materials and Methods-part 1

The experimental procedures for the isolation and characterisation of the *E. coli* O111 O antigen gene cluster (position 3,021-9,981) are according to Bastin D.A., et al. 1991 "Molecular cloning and expression in *Escherichia coli* K-12 of the *rfb* gene cluster determining the O antigen of an *E. coli* O111 strain". *Mol. Microbiol.* 5:9 2223-2231 and Bastin D.A. and Reeves, P.R. 1995 "Sequence and analysis of the O antigen gene(*rfb*) cluster of *Escherichia coli* O111". *Gene* 164: 17-23.

A. Bacterial strains and growth media

Bacteria were grown in Luria broth supplemented as required.

B. Cosmids and phage

Cosmids in the host strain x2819 were repackaged in vivo. Cells were grown in 250mL flasks containing 30mL of culture, with moderate shaking at 30°C to an optical density of 0.3 at 580 nm. The defective lambda prophage was induced by heating in a water bath at 45°C for 15min followed by an incubation at 37°C with vigorous shaking for 2hr. Cells were then lysed by the addition of 0.3mL chloroform and shaking for a further 10min. Cell debris were removed from 1mL of lysate by a 5min spin in a microcentrifuge, and the supernatant removed to a fresh microfuge tube. One drop of chloroform was added then shaken vigorously through the tube contents.

C. DNA preparation

Chromosomal DNA was prepared from bacteria grown overnight at 37°C in a volume of 30mL of Luria broth. After harvesting by centrifugation, cells were washed and

- 51 -

resuspended in 10mL of 50mMTris-HCl pH 8.0. EDTA was added and the mixture incubated for 20min. Then lysozyme was added and incubation continued for a further 10min. Proteinase K, SDS, and ribonuclease were then added and the mixture incubated for up to 2hr for lysis to occur. All incubations were at 37°C. The mixture was then heated to 65°C and extracted once with 8mL of phenol at the same temperature. The mixture was extracted once with 5mL of phenol/chloroform/iso-amyl alcohol at 4°C.

Residual phenol was removed by two ether extractions. DNA was precipitated with 2 vols. of ethanol at 4°C, spooled and washed in 70% ethanol, resuspended in 1-2mL of TE and dialysed. Plasmid and cosmid DNA was prepared by a modification of the Birnboim and Doly method [Birnboim, H. C. and Doly, J. (1979) "A rapid alkaline extraction procedure for screening recombinant plasmid DNA" *Nucl. Acid Res.* 7:1513-1523]. The volume of culture was 10mL and the lysate was extracted with phenol/chloroform/iso-amyl alcohol before precipitation with isopropanol. Plasmid DNA to be used as vector was isolated on a continuous caesium chloride gradient following alkaline lysis of cells grown in 1L of culture.

D. Enzymes and buffers.

Restriction endonucleases and DNA T4 ligase were purchased from Boehringer Mannheim (Castle Hill, NSW, Australia) or Pharmacia LKB (Melbourne, VIC Australia). Restriction enzymes were used in the recommended commercial buffer.

E. Construction of a gene bank.

Individual aliquots of M92 chromosomal DNA (strain Stoke W, from Statens Serum Institut, 5 Artillerivej, 2300 Copenhagen S, Denmark) were partially digested with 0.2U *Sau3A1* for 1-15mins. Aliquots giving the greatest proportion of fragments in the size range of approximately 40-50kb were selected and ligated to vector pPR691 previously digested with *Bam*H1 and *Pvu*II. Ligation mixtures were packaged *in vitro* with packaging extract.

- 52 -

The host strain for transduction was x2819 and recombinants were selected with kanamycin.

F. Serological procedures.

Colonies were screened for the presence of the O111 antigen by immunoblotting. Colonies were grown overnight, up to 100 per plate then transferred to nitrocellulose discs and lysed with 0.5N HCl. Tween 20 was added to TBS at 0.05% final concentration for blocking, incubating and washing steps. Primary antibody was *E. coli* O group 111 antiserum, diluted 1:800. The secondary antibody was goat anti-rabbit IgG labelled with horseradish peroxidase diluted 1:5000. The staining substrate was 4-chloro-1-naphthol. Slide agglutination was performed according to the standard procedure.

G. Recombinant DNA methods.

Restriction mapping was based on a combination of standard methods including single and double digests and sub-cloning. Deletion derivatives of entire cosmids were produced as follows: aliquots of 1.8mg of cosmid DNA were digested in a volume of 20ml with 0.25U of restriction enzyme for 5-80min. One half of each aliquot was used to check the degree of digestion on an agarose gel. The sample which appeared to give a representative range of fragments was ligated at 4°C overnight and transformed by the CaCl₂ method into JM109. Selected plasmids were transformed into sf174 by the same method. P4657 was transformed with pPR1244 by electroporation.

H. DNA hybridisation

Probe DNA was extracted from agarose gels by electroelution and was nick-translated using [α -³²P]-dCTP. Chromosomal or plasmid DNA was electrophoresed in 0.8% agarose and transferred to a nitrocellulose membrane. The hybridisation and pre-hybridisation buffers contained either 30% or 50% formamide for low and high stringency probing respectively. Incubation temperatures were 42°C and 37°C for pre-hybridisation and hybridisation respectively. Low stringency washing of

- 53 -

filters consisted of 3 x 20min washes in 2 x SSC and 0.1% SDS. High-stringency washing consisted of 3 x 5min washes in 2 x SSC and 0.1% SDS at room temperature, a 1hr wash in 1 x SSC and 0.1% SDS at 58°C and 15min wash in 0.1 x SSC and 0.1% SDS at 58°C.

I. Nucleotide sequencing of *E. coli* O111 O antigen gene cluster (position 3,021-9,981)

Nucleotide sequencing was performed using an ABI 373 automated sequencer (CA, USA). The region between map positions 3.30 and 7.90 was sequenced using uni-directional exonuclease III digestion of deletion families made in PT7T3190 from clones pPR1270 and pPR1272. Gaps were filled largely by cloning of selected fragments into M13mp18 or M13mp19. The region from map positions 7.90-10.2 was sequenced from restriction fragments in M13mp18 or M13mp19. Remaining gaps in both the regions were filled by priming from synthetic oligonucleotides complementary to determined positions along the sequence, using a single stranded DNA template in M13 or phagemid. The oligonucleotides were designed after analysing the adjacent sequence. All sequencing was performed by the chain termination method. Sequences were aligned using SAP [Staden, R., 1982 "Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing". *Nuc. Acid Res.* 10: 4731-4751; Staden, R., 1986 "The current status and portability of our sequence handling software". *Nuc. Acid Res.* 14: 217-231]. The program NIP [Staden, R. 1982 "An interactive graphics program for comparing and aligning nucleic acid and amino acid sequence". *Nuc. Acid Res.* 10: 2951-2961] was used to find open reading frames and translate them into proteins.

J. Isolation of clones carrying *E. coli* O111 O antigen gene cluster

The *E. coli* O antigen gene cluster was isolated according to the method of Bastin D.A., et al. [1991 "Molecular cloning and expression in *Escherichia coli* K-

12 of the *rfb* gene cluster determining the O antigen of an *E. coli* O111 strain". *Mol. Microbiol.* 5(9), 2223-2231]. Cosmid gene banks of M92 chromosomal DNA were established in the *in vivo* packaging strain x2819. From the genomic bank, 3.3×10^3 colonies were screened with *E. coli* O111 antiserum using an immuno-blotting procedure: 5 colonies (pPR1054, pPR1055, pPR1056, pPR1058 and pPR1287) were positive. The cosmids from these strains were packaged *in vivo* into lambda particles and transduced into the *E. coli* deletion mutant Sf174 which lacks all O antigen genes. In this host strain, all plasmids gave positive agglutination with O111 antiserum.

An *Eco* R1 restriction map of the 5 independent cosmids showed that they have a region of approximately 11.5 kb in common (Figure 1). Cosmid pPR1058 included sufficient flanking DNA to identify several chromosomal markers linked to O antigen gene cluster and was selected for analysis of the O antigen gene cluster region.

K. Restriction mapping of cosmid pPR1058

Cosmid pPR1058 was mapped in two stages. A preliminary map was constructed first, and then the region between map positions 0.00 and 23.10 was mapped in detail, since it was shown to be sufficient for O111 antigen expression. Restriction sites for both stages are shown in Figure 2. The region common to the five cosmid clones was between map positions 1.35 and 12.95 of pPR1058.

To locate the O antigen gene cluster within pPR1058, pPR1058 cosmid was probed with DNA probes covering O antigen gene cluster flanking regions from *S. enterica* LT2 and *E. coli* K-12. Capsular polysaccharide (*cps*) genes lie upstream of O antigen gene cluster while the gluconate dehydrogenase (*gnd*) gene and the histidine (*his*) operon are downstream, the latter being further from the O antigen gene cluster. The probes used were pPR472 (3.35kb), carrying the *gnd* gene of LT2, pPR685 (5.3kb) carrying two genes of the *cps* cluster, *cpsB* and

cpsG of LT2, and K350 (16.5kb) carrying all of the *his* operon of K-12. Probes hybridised as follows: pPR472 hybridised to 1.55kb and 3.5 kb (including 2.7 kb of vector) fragments of *Pst*I and *Hind*III double digests of pPR1246 (a *Hind*III/*Eco*R1 subclone derived from pPR1058, Figure 2), which could be located at map positions 12.95-15.1; pPR685 hybridised to a 4.4 kb *Eco*R1 fragment of pPR1058 (including 1.3 kb of vector) located at map position 0.00-3.05; and K350 hybridised with a 32kb *Eco*R1 fragment of pPR1058 (including 4.0kb of vector), located at map position 17.30-45.90. Subclones containing the presumed *gnd* region complemented a *gnd*⁻*edd*⁻ strain GB23152. On gluconate bromothymol blue plates, pPR1244 and pPR1292 in this host strain gave the green colonies expected of a *gnd*⁺*edd*⁻ genotype. The *his*⁺ phenotype was restored by plasmid pPR1058 in the *his* deletion strain Sf174 on minimal medium plates, showing that the plasmid carries the entire *his* operon.

It is likely that the O antigen gene cluster region lies between *gnd* and *cps*, as in other *E. coli* and *S. enterica* strains, and hence between the approximate map positions 3.05 and 12.95. To confirm this, deletion derivatives of pPR1058 were made as follows: first, pPR1058 was partially digested with *Hind*III and self ligated. Transformants were selected for kanamycin resistance and screened for expression of O111 antigen. Two colonies gave a positive reaction. *Eco*R1 digestion showed that the two colonies hosted identical plasmids, one of which was designated pPR1230, with an insert which extended from map positions 0.00 to 23.10. Second pPR1058 was digested with *Sal*I and partially digested with *Xho*I and the compatible ends were re-ligated. Transformants were selected with kanamycin and screened for O111 antigen expression. Plasmid DNA of 8 positively reacting clones was checked using *Eco*R1 and *Xho*I digestion and appeared to be identical. The cosmid of one was designated pPR1231. The insert of pPR1231

contained the DNA region between map positions 0.00 and 15.10. Third, pPR1231 was partially digested with *Xho*I, self-ligated, and transformants selected on spectinomycin/ streptomycin plates. Clones were screened for kanamycin sensitivity and of 10 selected, all had the DNA region from the *Xho*I site in the vector to the *Xho*I site at position 4.00 deleted. These clones did not express the O111 antigen, showing that the *Xho*I site at position 4.00 is within the O antigen gene cluster. One clone was selected and named pPR1288. Plasmids pPR1230, pPR1231, and pPR1288 are shown in Figure 2.

L. Analysis of the E. coli O111 O antigen gene cluster (position 3,021-9,981) nucleotide sequence data

Bastin and Reeves [1995 "Sequence and analysis of the O antigen gene(*rfb*)cluster of *Escherichia coli* O111". Gene 164: 17-23] partially characterised the *E. coli* O111 O antigen gene cluster by sequencing a fragment from map position 3,021-9,981. Figure 3 shows the gene organisation of position 3,021-9,981 of *E. coli* O111 O antigen gene cluster. *orf3* and *orf6* have high level amino acid identity with *wcaH* and *wcaG* (46.3% and 37.2% respectively), and are likely to be similar in function to sugar biosynthetic pathway genes in the *E. coli* K-12 colanic gene cluster. *orf4* and *orf5* show high levels of amino acid homology to *manC* and *manB* genes respectively. *orf7* shows high level homology with *rfbH* which is an abequose pathway gene. *orf8* encodes a protein with 12 transmembrane segments and has similarity in secondary structure to other *wzx* genes and is likely therefore to be the O antigen flippase gene.

Materials and Methods-part 2

A. Nucleotide sequencing of 1 to 3,020 and 9,982 to 14,516 of the *E. coli* O111 O antigen gene cluster

The sub clones which contained novel nucleotide sequences, pPR1231 (map position 0 and 1,510), pPR1237 (map position -300 to 2,744), pPR1239 (map position 2,744

to 4,168), pPR1245 (map position 9,736 to 12,007) and pPR1246 (map position 12,007 to 15,300) (Figure 2), were characterised as follows: the distal ends of the inserts of pPR1237, pPR1239 and pPR1245 were sequenced using the

M13 forward and reverse primers located in the vector. PCR walking was carried out to sequence further into each insert using primers based on the sequence data and the primers were tagged with M13 forward or reverse primer sequences for sequencing. This PCR walking procedure was repeated until the entire insert was sequenced. pPR1246 was characterised from position 12,007 to 14,516. The DNA of these sub clones was sequenced in both directions.

The sequencing reactions were performed using the dideoxy termination method and thermocycling and reaction products were analysed using fluorescent dye and an ABI automated sequencer (CA, USA).

B. Analysis of the *E. coli* O111 O antigen gene cluster (positions 1 to 3,020 and 9,982 to 14,516 of Figure 5) nucleotide sequence data

The gene organisation of regions of *E. coli* O111 O antigen gene cluster which were not characterised by Bastin and Reeves [1995 "Sequence and analysis of the O antigen gene(*rfb*)cluster of *Escherichia coli* O111." *Gene* 164: 17-23], (positions 1 to 3,020 and 9,982 to 14,516) is shown in Figure 3. There are two open reading frames in region 1. Four open reading frames are predicted in region 2. The position of each gene is listed in Table 9.

The deduced amino acid sequence of *orf1* (*wbdH*) shares about 64% similarity with that of the *rfp* gene of *Shigella dysenteriae*. Rfp and WbdH have very similar hydrophobicity plots and both have a very convincing predicted transmembrane segment in a corresponding position. *rfp* is a galactosyl transferase involved in the synthesis of LPS core, thus *wbdH* is likely to be a galactosyl transferase gene. *orf2* has 85.7% identity at amino acid level to the *gmd* gene identified in the *E.*

coli K-12 colanic acid gene cluster and is likely to be a *gmd* gene. *orf9* encodes a protein with 10 predicted transmembrane segments and a large cytoplasmic loop. This inner membrane topology is a characteristic feature of all known O antigen polymerases thus it is likely that *orf9* encodes an O antigen polymerase gene, *wzy*. *orf10* (*wbdL*) has a deduced amino acid sequence with low homology with *Lsi2* of *Neisseria gonorrhoeae*. *Lsi2* is responsible for adding GlcNAc to galactose in the synthesis of lipooligosaccharide. Thus it is likely that *wbdL* is either a colitose or glucose transferase gene. *orf11* (*wbdM*) shares high level nucleotide and amino acid similarity with *TrsE* of *Yersinia enterocolitica*. *TrsE* is a putative sugar transferase thus it is likely that *wbdM* encodes the colitose or glucose transferase.

In summary three putative transferase genes and an O antigen polymerase gene were identified at map position 1 to 3,020 and 9,982 to 14,516 of *E. coli* O111 O antigen gene cluster. A search of GenBank has shown that there are no genes with significant similarity at the nucleotide sequence level for two of the three putative transferase genes or the polymerase gene. Figure 5 provides the nucleotide sequence of the O111 antigen gene cluster.

Materials and Methods-part 3

A. PCR amplification of O157 antigen gene cluster from an *E. coli* O157:H7 strain (Strain C664-1992, from Statens Serum Institut, 5 Artillerivej, 2300, Copenhagen S, Denmark)

E. coli O157 O antigen gene cluster was amplified by using long PCR [Cheng et al. 1994, "Effective amplification of long targets from cloned inserts and human and genomic DNA" P.N.A.S. USA 91: 5695-569] with one primer (primer #412: att ggt agc tgt aag cca agg gcg gta gcg t) based on the JumpStart sequence usually found in the promoter region of O antigen gene clusters [Hobbs,

et al. 1994 "The JumpStart sequence: a 39 bp element common to several polysaccharide gene clusters" Mol. Microbiol. 12: 855-856], and another primer #482 (cac tgc cat acc gac gac gcc gat ctg ttg ctt gg) based on the *gnd* gene usually found downstream of the O antigen gene cluster. Long PCR was carried out using the Expand Long Template PCR System from Boehringer Mannheim (Castle Hill NSW Australia), and products, 14 kb in length, from several reactions were combined and purified using the Promega Wizard PCR preps DNA purification System (Madison WI USA). The PCR product was then extracted with phenol and twice with ether, precipitated with 70% ethanol, and resuspended in 40mL of water.

B. Construction of a random DNase I bank:

Two aliquots containing about 150ng of DNA each were subjected to DNase I digestion using the Novagen DNase I Shotgun Cleavage (Madison WI USA) with a modified protocol as described. Each aliquot was diluted into 45ml of 0.05M Tris -HCl (pH7.5), 0.05mg/mL BSA and 10mM MnCl₂. 5mL of 1:3000 or 1:4500 dilution of DNaseI (Novagen) (Madison WI USA) in the same buffer was added into each tube respectively and 10ml of stop buffer (100mM EDTA), 30% glycerol, 0.5% Orange G, 0.075% xylene and cyanol (Novagen) (Madison WI USA) was added after incubation at 15°C for 5 min. The DNA from the two DNaseI reaction tubes were then combined and fractionated on a 0.8% LMT agarose gel, and the gel segment with DNA of about 1kb in size (about 1.5mL agarose) was excised. DNA was extracted from agarose using Promega Wizard PCR Preps DNA Purification (Madison WI USA) and resuspended in 200 mL water, before being extracted with phenol and twice with ether, and precipitated. The DNA was then resuspended in 17.25 mL water and subjected to T4 DNA polymerase repair and single dA tailing using the Novagen Single dA Tailing Kit (Madison WI USA). The reaction product (85ml containing about 8ng DNA) was then extracted with chloroform:isoamyl alcohol (24:1) once and

ligated to 3×10^{-3} pmol pGEM-T (Promega) (Madison WI USA) in a total volume of 100mL. Ligation was carried out overnight at 4°C and the ligated DNA was precipitated and resuspended in 20mL water before being electroporated into *E. coli* strain JM109 and plated out on BCIG-IPTG plates to give a bank.

C. Sequencing

DNA templates from clones of the bank were prepared for sequencing using the 96-well format plasmid DNA miniprep kit from Advanced Genetic Technologies Corp (Gaithersburg MD USA). The inserts of these clones were sequenced from one or both ends using the standard M13 sequencing primer sites located in the pGEM-T vector. Sequencing was carried out on an ABI377 automated sequencer (CA USA) as described above, after carrying out the sequencing reaction on an ABI Catalyst (CA USA). Sequence gaps and areas of inadequate coverage were PCR amplified directly from O157 chromosomal DNA using primers based on the already obtained sequencing data and sequenced using the standard M13 sequencing primer sites attached to the PCR primers.

D. Analysis of the *E. coli* O157 O antigen gene cluster nucleotide sequence data

Sequence data were processed and analysed using the Staden programs [Staden, R., 1982 "Automation of the computer handling of gel reading data produced by the shotgun method of DNA sequencing." *Nuc. Acid Res.* 10: 4731-4751; Staden, R., 1986 "The current status and portability of our sequence handling software". *Nuc. Acid Res.* 14: 217-231; Staden, R. 1982 "An interactive graphics program for comparing and aligning nucleic acid and amino acid sequence". *Nuc. Acid Res.* 10: 2951-2961].

Figure 4 shows the structure of *E. coli* O157 O antigen gene cluster. Twelve open reading frames were predicted from the sequence data, and the nucleotide and amino acid sequences of all these genes were then used to search the GenBank database for indication of possible function and

specificity of these genes. The position of each gene is listed in Table 9. The nucleotide sequence is presented in Figure 6.

5 orfs 10 and 11 showed high level identity to *manC* and *manB* and were named *manC* and *manB* respectively. *orf7* showed 89% identity (at amino acid level) to the *gmd* gene of the *E. coli* colanic acid capsule gene cluster (Stevenson G., K. et al. 1996 "Organisation of the *Escherichia coli* K-12 gene cluster responsible for production of the extracellular polysaccharide colanic acid". J. Bacteriol. 178:4885-4893) and was named *gmd*.
10 *orf8* showed 79% and 69% identity (at amino acid level) respectively to *wcaG* of the *E. coli* colanic acid capsule gene cluster and to *wbcJ* (*orf14.8*) gene of the *Yersinia enterocolitica* O8 O antigen gene cluster (Zhang, L. et al. 1997 "Molecular and chemical characterization of the lipopolysaccharide O-antigen and its role in the virulence of *Y. enterocolitica* serotype O8". Mol. Microbiol. 23:63-76). Colanic acid and the *Yersinia* O8 O
15 antigen both contain fucose as does the O157 O antigen. There are two enzymatic steps required for GDP-L-fucose synthesis from GDP-4-keto-6-deoxy-D-mannose, the product of the *gmd* gene product. However, it has been shown recently (Tonetti, M et al. 1996 Synthesis of GDP-L-fucose by the human FX protein J. Biol. Chem. 271:27274-27279) that the human FX protein has "significant
20 homology" with the *wcaG* gene (referred to as *Yefb* in that paper), and that the FX protein carries out both reactions to convert GDP-4-keto-6-deoxy-D-mannose to GDP-L-fucose. We believe that this makes a very strong case for *orf8* carrying out these two steps and propose to name the gene *fcl*. In support of the one enzyme carrying out both functions is the observation that there are no genes other than *manB*, *manC*, *gmd* and *fcl* with similar levels of
25 similarity between the three bacterial gene clusters for fucose containing structures.

35 *orf5* is very similar to *wbeE* (*rfbE*) of *Vibrio*

cholerae O1, which is thought to be the perosamine synthetase, which converts GDP-4-keto-6-deoxy-D-mannose to GDP-perosamine (Stroeher, U.H et al. 1995 "A putative pathway for perosamine biosynthesis is the first function encoded within the *rfb* region of *Vibrio cholerae*" O1. Gene 166: 33-42). *V. cholerae* O1 and *E. coli* O157 O antigens contain perosamine and N-acetyl-perosamine respectively. The *V. cholerae* O1 *manA*, *manB*, *gmd* and *wbeE* genes are the only genes of the *V. cholerae* O1 gene cluster with significant similarity to genes of the *E. coli* O157 gene cluster and we believe that our observations both confirm the prediction made for the function of *wbe* of *V. cholerae*, and show that *orf5* of the O157 gene cluster encodes GDP-perosamine synthetase.

orf5 is therefore named *per*. *orf5* plus about 100bp of the upstream region (position 4022-5308) was previously sequenced by Bilge, S.S. et al. [1996 "Role of the *Escherichia coli* O157-H7 O side chain in adherence and analysis of an *rfb* locus". Infect. Immun. 64:4795-4801].

orf12 shows high level similarity to the conserved region of about 50 amino acids of various members of an acetyltransferase family (Lin, W., et al. 1994 "Sequence analysis and molecular characterisation of genes required for the biosynthesis of type 1 capsular polysaccharide in *Staphylococcus aureus*". J. Bacteriol. 176: 7005-7016) and we believe it is the N-acetyltransferase to convert GDP-perosamine to GDP-perNAC. *orf12* has been named *wbdR*.

The genes *manB*, *manC*, *gmd*, *fcl*, *per* and *wbdR* account for all of the expected biosynthetic pathway genes of the O157 gene cluster.

The remaining biosynthetic step(s) required are for synthesis of UDP-GalNAC from UDP-Glc. It has been proposed (Zhang, L., et al. 1997 "Molecular and chemical characterisation of the lipopolysaccharide O-antigen and its role in the virulence of *Yersinia enterocolitica* serotype O8". Mol. Microbiol. 23:63-76) that in *Yersinia enterocolitica* UDP-GalNAC is synthesised from UDP-GlcNAC

by a homologue of galactose epimerase (GalE), for which there is a *galE* like gene in the *Yersinia enterocolitica* O8 gene cluster. In the case of O157 there is no *galE* homologue in the gene cluster and it is not clear how UDP-GalNAc is synthesised. It is possible that the galactose epimerase encoded by the *galE* gene in the *gal* operon, can carry out conversion of UDP-GlcNAc to UDP-GalNAc in addition to conversion of UDP-Glc to UDP-Gal. There do not appear to be any gene(s) responsible for UDP-GalNAc synthesis in the O157 gene cluster.

orf4 shows similarity to many *wzx* genes and is named *wzx* and *orf2* which shows similarity of secondary structure in the predicted protein to other *wzy* genes and is for that reason named *wzy*.

The *orf1*, *orf3* and *orf6* gene products all have characteristics of transferases, and have been named *wbdN*, *wbdO* and *wbdP* respectively. The O157 O antigen has 4 sugars and 4 transferases are expected. The first transferase to act would put a sugar phosphate onto undecaprenol phosphate. The two transferases known to perform this function, WbaP (RfbP) and WecA (Rfe) transfer galactose phosphate and N-acetyl-glucosamine phosphate respectively to undecaprenol phosphate. Neither of these sugars is present in the O157 structure.

Further, none of the presumptive transferases in the O157 gene cluster has the transmembrane segments found in WecA and WbaP which transfer a sugar phosphate to undecaprenol phosphate and expected for any protein which transferred a sugar to undecaprenol phosphate which is embedded within the membrane.

The WecA gene which transfers GlcNAc-P to undecaprenol phosphate is located in the Enterobactereal Common Antigen (ECA) gene cluster and it functions in ECA synthesis in most and perhaps all *E. coli* strains, and also in O antigen synthesis for those strains which have GlcNAc as the first sugar in the O unit.

It appears that WecA acts as the transferase for

- 64 -

addition of GalNAc-1-P to undecaprenol phosphate for the *Yersinia enterocolitica* O8 O antigen [Zhang et al.1997 "Molecular and chemical characterisation of the lipopolysaccharide O antigen and its role in the virulence of *Yersinia enterocolitica* serotype O8" Mol. Microbiol. 23: 63-76.] and perhaps does so here as the O157 structure includes GalNAc. *WecA* has also been reported to add Glucose-1-P phosphate to undecaprenol phosphate in *E. coli* O8 and O9 strains, and an alternative possibility for transfer of the first sugar to undecaprenol phosphate is *WecA* mediated transfer of glucose, as there is a glucose residue in the O157 O antigen. In either case the requisite number of transferase genes are present if GalNAc or Glc is transferred by *WecA* and the side chain Glc is transferred by a transferase outside of the O antigen gene cluster.

orf9 shows high level similarity (44% identity at amino acid level, same length) with *wcaH* gene of the *E. coli* colanic acid capsule gene cluster. The function of this gene is unknown, and we give *orf9* the name *wbdQ*.

The DNA between *manB* and *wdbR* has strong sequence similarity to one of the H-repeat units of *E. coli* K12. Both of the inverted repeat sequences flanking this region are still recognisable, each with two of the 11 bases being changed. The H-repeat associated protein encoding gene located within this region has a 267 base deletion and mutations in various positions. It seems that the H-repeat unit has been associated with this gene cluster for a long period of time since it translocated to the gene cluster, perhaps playing a role in assembly of the gene cluster as has been proposed in other cases.

Materials and Methods - part 4

To test our hypothesis that O antigen genes for transferases and the *wzx*, *wzy* genes were more specific than pathway genes for diagnostic PCR, we first carried out PCR using primers for all the *E. coli* O16 O antigen

genes (Table 7). The PCR was then carried out using PCR primers for *E. coli* 0111 transferase, *wzx* and *wzy* genes (Table 8, 8A). PCR was also carried out using PCR primers for the *E. coli* 0157 transferase, *wzx* and *wzy* genes (Table 9, 9A).

Chromosomal DNA from the 166 serotypes of *E. coli* available from Statens Serum Institut, 5 Artillerivej, 2300 Copenhagen Denmark was isolated using the Promega Genomic (Madison WI USA) isolation kit. Note that 164 of the serogroups are described by Ewing W. H.: Edwards and Ewings "Identification of the Enterobacteriaceae" Elsevier, Amsterdam 1986 and that they are numbered 1-171 with numbers 31, 47, 67, 72, 93, 94 and 122 no longer valid. Of the two serogroup 19 strains we used 19ab strain F8188-41. Lior H. 1994 ["Classification of *Escherichia coli* In *Escherichia coli* in domestic animals and humans pp 31-72. Edited by C.L. Gyles CAB international] adds two more numbered 172 and 173 to give the 166 serogroups used. Pools containing 5 to 8 samples of DNA per pool were made. Pool numbers 1 to 19 (Table 4) were used in the *E. coli* 0111 and 0157 assay. Pool numbers 20 to 28 were also used in the 0111 assay, and pool numbers 22 to 24 contained *E. coli* 0111 DNA and were used as positive controls (Table 5). Pool numbers 29 to 42 were also used in the 0157 assay, and pool numbers 31 to 36 contained *E. coli* 0157 DNA, and were used as positive controls (Table 6). Pool numbers 2 to 20, 30, 43 and 44 were used in the *E. coli* 016 assay (Tables 4 to 6). Pool number 44 contained DNA of *E. coli* K-12 strains C600 and WG1 and was used as a positive control as between them they have all of the *E. coli* K-12 016 O antigen genes.

PCR reactions were carried out under the following conditions: denaturing 94°C/30"; annealing, temperature varies (refer to Tables)/30"; extension, 72°C/1'; 30 cycles. PCR reaction was carried out in a volume of 25mL for each pool. After the PCR reaction, 10mL PCR

product from each pool was run on an agarose gel to check for amplified DNA.

Each *E. coli* chromosomal DNA sample was checked by gel electrophoresis for the presence of chromosomal DNA and by PCR amplification of the *E. coli mdh* gene using oligonucleotides based on *E. coli* K-12 [Boyd et al. (1994) "Molecular genetic basis of allelic polymorphism in malate dehydrogenase (*mdh*) in natural populations of *Escherichia coli* and *Salmonella enterica*" Proc. Nat. Acad. Sci. USA. 91:1280-1284.] Chromosomal DNA samples from other bacteria were only checked by gel electrophoresis of chromosomal DNA.

A. Primers based on *E. coli* O16 O antigen gene cluster sequence.

The O antigen gene cluster of *E. coli* O16 was the only typical *E. coli* O antigen gene cluster that had been fully sequenced prior to that of O111, and we chose it for testing our hypothesis. One pair of primers for each gene was tested against pools 2 to 20, 30 and 43 of *E. coli* chromosomal DNA. The primers, annealing temperatures and functional information for each gene are listed in Table 8.

For the five pathway genes, there were 17/21, 13/21, 0/21, 0/21, 0/21 positive pools for *rmlB*, *rmlD*, *rmlA*, *rmlC* and *glf* respectively (Table 7). For the *wzx*, *wzy* and three transferase genes there were no positives amongst the 21 pools of *E. coli* chromosomal DNA tested (Table 7). In each case the #44 pool gave a positive result.

B. Primers based on the *E. coli* O111 O antigen gene cluster sequence.

One to four pairs of primers for each of the transferase, *wzx* and *wzy* genes of O111 were tested against the pools 1 to 21 of *E. coli* chromosomal DNA (Table 8). For *wbdH*, four pairs of primers, which bind

to various regions of this gene, were tested and found to be specific for O111 as there was no amplified DNA of the correct size in any of those 21 pools of *E. coli* chromosomal DNA tested. Three pairs of primers for *wbdM* were tested, and they are all specific although primers #985/#986 produced a band of the wrong size from one pool. Three pairs of primers for *wzx* were tested and they all were specific. Two pairs of primers were tested for *wzy*, both are specific although #980/#983 gave a band of the wrong size in all pools. One pair of primers for *wbdL* was tested and found unspecific and therefore no further test was carried out. Thus, *wzx*, *wzy* and two of the three transferase genes are highly specific to O111.

Bands of the wrong size found in amplified DNA are assumed to be due to chance hybridisation of genes widely present in *E. coli*. The primers, annealing temperatures and positions for each gene are in Table 8.

The O111 assay was also performed using pools including DNA from O antigen expressing *Yersinia pseudotuberculosis*, *Shigella boydii* and *Salmonella enterica* strains (Table 8A). None of the oligonucleotides derived from *wbdH*, *wzx*, *wzy* or *wbdM* gave amplified DNA of the correct size with these pools. Notably, pool number 25 includes *S. enterica* Adelaide which has the same O antigen as *E. coli* O111: this pool did not give a positive PCR result for any primers tested indicating that these genes are highly specific for *E. coli* O111.

Each of the 12 pairs binding to *wbdH*, *wzx*, *wzy* and *wbdM* produces a band of predicted size with the pools containing O111 DNA (pools number 22 to 24). As pools 22 to 24 included DNA from all strains present in pool 21 plus O111 strain DNA (Table 5), we conclude that the 12 pairs of primers all give a positive PCR test with each of three unrelated O111 strains but not with any other strains tested. Thus these genes are highly specific for *E. coli* O111.

C. Primers based on the *E. coli* 0157 O antigen gene cluster sequence.

Two or three primer pairs for each of the transferase, *wzx* and *wzy* genes of 0157 were tested against *E. coli* chromosomal DNA of pools 1 to 19, 29 and 30 (Table 9). For *wbdN*, three pairs of primers, which bind to various regions of this gene, were tested and found to be specific for 0157 as there was no amplified DNA in any of those 21 pools of *E. coli* chromosomal DNA tested. Three pairs of primers for *wbdO* were tested, and they are all specific although primers # 1211/#1212 produced two or three bands of the wrong size from all pools. Three pairs of primers were tested for *wbdP* and they all were specific. Two pairs of primers were tested for *wbdR* and they were all specific. For *wzy*, three pairs of primers were tested and all were specific although primer pair #1203/#1204 produced one or three bands of the wrong size in each pool. For *wzx*, two pairs of primers were tested and both were specific although primer pair #1217/#1218 produced 2 bands of wrong size in 2 pools, and 1 band of wrong size in 7 pools. Bands of the wrong size found in amplified DNA are assumed to be due to chance hybridisation of genes widely present in *E. coli*. The primers, annealing temperatures and function information for each gene are in Table 9.

The 0157 assay was also performed using pools 37 to 42, including DNA from O antigen expressing *Yersinia pseudotuberculosis*, *Shigella boydii*, *Yersinia enterocolitica* 09, *Brucella abortus* and *Salmonella enterica* strains (Table 9A). None of the oligonucleotides derived from *wbdN*, *wzy*, *wbdO*, *wzx*, *wbdP* or *wbdR* reacted specifically with these pools, except that primer pair #1203/#1204 produced two bands with *Y. enterocolitica* 09 and one of the bands is of the same size with that from the positive control. Primer pair #1203/#1204 binds to *wzy*. The predicted secondary

- 69 -

structures of Wzy proteins are generally similar, although there is very low similarity at amino acid or DNA level among the sequenced wzy genes. Thus, it is possible that *Y. enterocolitica* 09 has a wzy gene closely related to that of *E. coli* 0157. It is also possible that this band is due to chance hybridization of another gene, as the other two wzy primer pairs (#1205/#1206 and #1207/#1208) did not produce any band with *Y. enterocolitica* 09. Notably, pool number 37 includes *S. enterica* Landau which has the same O antigen as *E. coli* 0157, and pool 38 and 39 contain DNA of *B. abortus* and *Y. enterocolitica* 09 which cross react serologically with *E. coli* 0157. This result indicates that these genes are highly 0157 specific, although one primer pair may have cross reacted with *Y. enterocolitica* 09.

Each of the 16 pairs binding to *wbdN*, *wzx*, *wzy*, *wbdO*, *wbdP* and *wbdR* produces a band of predicted size with the pools containing 0157 DNA (pools number 31 to 36). As pool 29 included DNA from all strains present in pools 31 to 36 other than 0157 strain DNA (Table 6), we conclude that the 16 pairs of primers all give a positive PCR test with each of the five unrelated 0157 strains.

Thus PCR using primers based on genes *wbdN*, *wzy*, *wbdO*, *wzx*, *wbdP* and *wbdR* is highly specific for *E. coli* 0157, giving positive results with each of six unrelated 0157 strains while only one primer pair gave a band of the expected size with one of three strains with O antigens known to cross-react serologically with *E. coli* 0157.

TABLE 1

H7 strains used in this work in addition to the H
antigens type strains

5

Name used in this study	Serotype	Original name	Source*
M527	O157:H7	C664-1992	a
M917	O18ac:H7	A57	IMVS
M918	O18ac:H7	A62	IMVS
M973	O2:H7	A1107	CDC
M1004	O157:H7	EH7	b
M1179	O18ac:H7	D-M3291/54	IMVS
M1200	O7:H7	A64	c
M1211	O19ab:H7	F8188-41	IMVS
M1328	O53:H7	14097	IMVS
M1686	O55:H7	TB156	d

*

- 10 a. Statens Serum Institut, Copenhagen, Denmark.
- b. Dr R. Brown of Royal Children's Hospital, Melbourne, Australia.
- 15 c. Max-Planck Institut fur molekulare Genetik, Berlin, Germany.
- d. Dr P. Tarr of Children's Hospital and Medical Center, University of Washington, USA.
- 20 IMVS, Institute of Medical and veterinary Science, Adelaide, Australia.
- CDC, Centers for Disease Control and prevention, Atlanta, USA.

Table 2

Oligonucleotides used to PCR amplify *fliC* genes
from different H type strains for sequencing

H Type Strains	Annealing Temperature (°C)	Primers Used
1	55	#1575/#1576
2	55	#1285/#1286
3	55	#1285/#1286
4	50	#1431/#1432
5	60	#1285/#1286
6	55	#1575/#1576
7	55	#1575/#1576
8	55	#1431/#1432
9	60	#1575/#1576
10	55	#1575/#1576
11	55	#1285/#1286
12	60	#1575/#1576
14	60	#1575/#1576
15	60	#1575/#1576
16	60	#1575/#1576
17	60	#1417/#1418
18	60	#1575/#1576
19	60	#1575/#1576
20	60	#1575/#1576
21	55	#1285/#1286
23	60	#1575/#1576
24	60	#1285/#1286
25	60	#1417/#1418
26	60	#1575/#1576
27	50	#1431/#1432
28	60	#1575/#1576
29	60	#1285/#1286
30	60	#1575/#1576
31	60	#1575/#1576
32	60	#1575/#1576
33	60	#1285/#1286
34	55	#1575/#1576
35	50	#1431/#1432
37	60	#1285/#1286
38	60	#1285/#1286
39	55	#1285/#1286
40	55	#1285/#1286
41	60	#1575/#1576
42	60	#1285/#1286
43	60	#1575/#1576
44	60	#1285/#1286
45	60	#1575/#1576
46	60	#1575/#1576
47	55	#1285/#1286
48	60	#1575/#1576
49	60	#1575/#1576
50	60	#1285/#1286
51	60	#1575/#1576
52	60	#1575/#1576
54	50	#1431/#1432
55	60	#1285/#1286
56	60	#1285/#1286

Table 3
Summary of the flagellin sequences obtained and specific H type
oligonucleotide primers

H type strain(s) the sequenced gene(s) obtained from	H specificity coded by the gene(s)	H type strain from which the flagellin gene sequence was used for primer choice	Positions of primer 1	Positions of primer 2
1	1	1	892-909	1172-1189
2	2	2	568-587	1039-1056
4,17,44	4	4	466-483	628-648
5	5	5	697-714	877-897
6	6	6	565-585	799-816
7	7	7	553-570 (primer #1806)	1483-1500 (primer #1809)
9	9	9	616-633	838-855
10(50)***	10	10	559-579	697-717
11	11	11	586-606*	791-810*
12	12	12	892-909	1172-1189
14	14	14	586-606	793-813
15	15	15	640-660	817-834
3	16	3	649-666	925-942
18	18	18	589-606	802-819
19	19	19	607-624	538-855
20	20	20	574-591	760-780
21,47	21	21	676-693**	862-879**
23	23	23	637-654	1336-1353
24	24	24	496-516	772-792
26	26	26	553-570	772-789
27	27	27	685-702	799-819
28	28	28	592-609	778-798
29	29	29	538-555	757-774
30	30	30	814-831	943-962
31	31	31	571-588	790-807
32	32	32	514-831	1057-1074
33	33	33	553-570	718-735
34	34	34	568-585	796-816
38,55	38	38	553-573	709-729
39	39	39	556-573	718-735
41	41	41	598-615	784-801
42	42	42	547-567	715-735
43	43	43	580-597	844-861
45	45	45	640-657	943-963
46	46	46	565-582	781-801
49	49	49	589-609	754-771
51	51	51	565-582	1042-1059
52	52	52	598-615	829-846
56	56	56	697-714	877-897
8 and 40		8	562-579	1045-1062
25		25	529-549	703-723
35		non-functional H11 gene	769-789*	1045-1065*
37		37	520-537	715-735
48		48	568-585	835-852
54		non-functional H21 gene	988-1008**	1344-1364**

* See section 13 for choice of primers for the flagellin gene of H11

** See section 13 for choice of primers for the flagellin gene of H21

*** See text

Table 3A
Cloning, expression and identification of flagellin genes

H type strain from which the H antigen gene was amplified	Primers used for PCR amplification of the H antigen gene	Annealing temperature (°C) used for PCR amplification	Plasmid carrying the H antigen gene	Host strain used for expression	Anti-serum which reacts with an <i>E. Coli</i> <i>fliC</i> deletion strain carrying the plasmid	H antigen encoded by the cloned gene
H1	#1868 & #1870	55	pPR1920	M2126	H1	H1
H2	#1868 & #1870	55	pPR1977	P5560	H2	H2
H3	#1868 & #1870	55	pPR1969	P5560	H16	H16
H4	#1878 & #1885	65	pPR1955	P5560	H4	H4
H5	#1868 & #1870	60	pPR1967	M2126	H5	H5
H6	#1868 & #1870	55	pPR1921	P5560	H6	H6
H7	#1868 & #1870	55	pPR1919	P5560	H7	H7
H9	#1868 & #1870	55	pPR1922	P5560	H9	H9
H10	#1868 & #1870	55	pPR1923	P5560	H10	H10
H11	#1868 & #1870	55	pPR1981	M2126	H11	H11
H12	#1868 & #1870	60	pPR1990	M2126	H12	H12
H14	#1868 & #1870	55	pPR1924	P5560	H14	H14
H15	#1868 & #1870	55	pPR1925	P5560	H15	H15
H17	#1878 & #1885	65	pPR1957	P5560	H4	H4
H18	#1868 & #1870	55	pPR1986	M2126	H18	H18
H19	#1868 & #1870	55	pPR1927	P5560	H19	H19
H20	#1868 & #1870	55	pPR1963	M2126	H20	H20
H21	#1868 & #1870	55	pPR1995	M2126	H21	H21
H23	#1868 & #1869	55	pPR1942	P5560	H23	H23
H24	#1868 & #1870	55	pPR1971	M2126	H24	H24
H26	#1868 & #1870	65	pPR1928	P5560	H26	H26
H27	#1868 & #1870	55	pPR1970	M2126	H27	H27
H28	#1868 & #1870	60	pPR1944	P5560	H28	H28
H29	#1868 & #1870	55	pPR1972	M2126	H29	H29
H30	#1868 & #1871	55	pPR1948	P5560	H30	H30
H31	#1868 & #1870	65	pPR1965	M2126	H31	H31
H32	#1868 & #1871	55	pPR1940	P5560	H32	H32
H33	#1868 & #1871	55	pPR1976	M2126	H33	H33
H34	#1868 & #1870	65	pPR1930	P5560	H34	H34
H38	#1868 & #1870	48	pPR1984	M2126	H38	H38
H39	#1868 & #1870	48	pPR1982	M2126	H39	H39
H41	#1868 & #1870	65	pPR1931	P5560	H41	H41
H42	#1868 & #1870	50	pPR1979	M2126	H42	H42
H43	#1868 & #1870	65	pPR1968	M2126	H43	H43
H45	#1868 & #1870	60	pPR1943	P5560	H45	H45
H46	#1868 & #1870	60	pPR1966	M2126	H46	H46
H49	#1868 & #1870	60	pPR1985	M2126	H49	H49
H51	#1868 & #1870	65	pPR1941	P5560	H51	H51
H52	#1868 & #1870	65	pPR1935	P5560	H52	H52
H56	#1868 & #1870	50	pPR1978	M2126	H56	H56

Table 3B Oligonucleotide primers used for PCR amplification and cloning of H antigen genes

#1868 5'- cat gcc atg gca caa gtc att aat acc -3'

NcoI

#1869 5'- ata tgt cga ctt aac cct gca gca gag aca g -3'

SalI

#1870 5' - atg gat cct taa ccc tgc agc aga gac ag -3'

BamHI

#1871 5' - aac tgc agt taa ccc tgt agc aga gac ag -3'

PstI

#1872 5' - cgg gat ccc gca gac tgg ttc ttg ttg at - 3'

BamHI

#1878 5' - cgg gat cca ctt cta tcg agc gcc tct ct - 3'

BamHI

#1884 5' - gct cta gag cgc aga tca ttc agc agg cc -3'

XbaI

#1885 5' - gct cta gac atg ttg gac act tcg gtc gc - 3'

XbaI

- 75 -

TABLE 4

Pool No.	Strains of which chromosomal DNA included in the pool	Source*
1	<i>E. coli</i> type strains for O serotypes 1, 2, 3, 4, 10, 16, 18 and 39	IMVS ^a
2	<i>E. coli</i> type strains for O serotypes 40, 41, 48, 49, 71, 73, 88 and 100	IMVS
3	<i>E. coli</i> type strains for O serotypes 102, 109, 119, 120, 121, 125, 126 and 137	IMVS
4	<i>E. coli</i> type strains for O serotypes 138, 139, 149, 7, 5, 6, 11 and 12	IMVS
5	<i>E. coli</i> type strains for O serotypes 13, 14, 15, 17, 19ab, 20, 21 and 22	IMVS
6	<i>E. coli</i> type strains for O serotypes 23, 24, 25, 26, 27, 28, 29 and 30	IMVS
7	<i>E. coli</i> type strains for O serotypes 32, 33, 34, 35, 36, 37, 38 and 42	IMVS
8	<i>E. coli</i> type strains for O serotypes 43, 44, 45, 46, 50, 51, 52 and 53	IMVS
9	<i>E. coli</i> type strains for O serotypes 54, 55, 56, 57, 58, 59, 60 and 61	IMVS
10	<i>E. coli</i> type strains for O serotypes 62, 63, 64, 65, 66, 68, 69 and 70	IMVS
11	<i>E. coli</i> type strains for O serotypes 74, 75, 76, 77, 78, 79, 80 and 81	IMVS
12	<i>E. coli</i> type strains for O serotypes 82, 83, 84, 85, 86, 87, 89 and 90	IMVS
13	<i>E. coli</i> type strains for O serotypes 91, 92, 95, 96, 97, 98, 99 and 101	IMVS
14	<i>E. coli</i> type strains for O serotypes 103, 104, 105, 106, 107, 108 and 110	IMVS
15	<i>E. coli</i> type strains for O serotypes 112, 162, 113, 114, 115, 116, 117 and 118	IMVS
16	<i>E. coli</i> type strains for O serotypes 123, 165, 166, 167, 168, 169, 170 and 171	See b
17	<i>E. coli</i> type strains for O serotypes 172, 173, 127, 128, 129, 130, 131 and 132	See c
18	<i>E. coli</i> type strains for O serotypes 133, 134, 135, 136, 140, 141, 142 and 143	IMVS
19	<i>E. coli</i> type strains for O serotypes 144, 145, 146, 147, 148, 150, 151 and 152	IMVS

*

- a. Institute of Medical and Veterinary Science, Adelaide, Australia
- b. 123 from IMVS; the rest from Statens Serum Institut, Copenhagen, Denmark
- c. 172 and 173 from Statens Serum Institut, Copenhagen, Denmark, the rest from IMVS

TABLE 5

Pool No.	Strains of which chromosomal DNA included in the pool	Source*
20	<i>E. coli</i> type strains for O serotypes 153, 154, 155, 156, 157, 158, 159 and 160	IMVS
21	<i>E. coli</i> type strains for O serotypes 161, 163, 164, 8, 9 and 124	IMVS
22	As pool #21, plus <i>E. coli</i> 0111 type strain Stoke W.	IMVS
23	As pool #21, plus <i>E. coli</i> 0111:H2 strain C1250-1991	See d
24	As pool #21, plus <i>E. coli</i> 0111:H12 strain C156-1989	See e
25	As pool #21, plus <i>S. enterica</i> serovar Adelaide	See f
26	<i>Y. pseudotuberculosis</i> strains of O groups IA, IIA, IIB, IIC, III, IVA, IVB, VA, VB, VI and VII	See g
27	<i>S. boydii</i> strains of serogroups 1, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14 and 15	See h
28	<i>S. enterica</i> strains of serovars (each representing a different O group) Typhi, Montevideo, Ferruch, Jangwani, Raus, Hvittingfoss, Waycross, Dan, Dugbe, Basel, 65:i:e,n,z,15 and 52:d:e,n,x,z15	IMVS

*

- d. C1250-1991 from Statens Serum Institut, Copenhagen, Denmark
- e. C156-1989 from Statens Serum Institut, Copenhagen, Denmark
- f. *S. enterica* serovar Adelaide from IMVS
- g. Dr S Aleksic of Institute of Hygiene, Germany
- h. Dr J Lefebvre of Bacterial Identification Section, Laboratoire de Santé Publique du Québec, Canada

TABLE 6

Pool No.	Strains of which chromosomal DNA included in the pool	Source*
29	<i>E. coli</i> type strains for O serotypes 153, 154, 155, 156, 158, 159 and 160	IMVS
30	<i>E. coli</i> type strains for O serotypes 161, 163, 164, 8, 9, 111 and 124	IMVS
31	As pool #29, plus <i>E. coli</i> O157 type strain A2 (O157:H19)	IMVS
32	As pool #29, plus <i>E. coli</i> O157:H16 strain C475-89	See d
33	As pool #29, plus <i>E. coli</i> O157:H45 strain C727-89	See d
34	As pool #29, plus <i>E. coli</i> O157:H2 strain C252-94	See d
35	As pool #29, plus <i>E. coli</i> O157:H39 strain C258-94	See d
36	As pool #29, plus <i>E. coli</i> O157:H26	See e
37	As pool #29, plus <i>S. enterica</i> serovar Landau	See f
38	As pool #29, plus <i>Brucella abortus</i>	See g See h
39	As pool #29, plus <i>Y. enterocolitica</i> O9	
40	<i>Y. pseudotuberculosis</i> strains of O groups IA, IIA, IIB, IIC, III, IVA, IVB, VA, VB, VI and VII	See i
41	<i>S. boydii</i> strains of serogroups 1, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14 and 15	See j
42	<i>S. enterica</i> strains of serovars (each representing a different O group) Typhi, Montevideo, Ferruch, Jangwani, Raus, Hvittingfoss, Waycross, Dan, Dugbe, Basel, 65:i:e,n,z15 and 52:d:e,n,x,z15	IMVS
43	<i>E. coli</i> type strains for O serotypes 1,2,3,4,10,18 and 29	IMVS
44	As pool #43, plus <i>E. coli</i> K-12 strains C600 and WG1	IVMS See k

*

- d. O157 strains from Statens Serum Institut, Copenhagen, Denmark
- e. O157:H26 from Dr R Brown of Royal Children's Hospital, Melbourne, Victoria
- f. *S. enterica* serovar Landau from Dr M Poppoff of Institut Pasteur, Paris, France
- g. *B. Abortus* from the culture collection of The University of Sydney, Sydney, Australia
- h. *Y. enterocolitica* O9 from Dr. K. Bettelheim of Victorian Infectious Diseases Reference Laboratory Victoria, Australia.
- i. Dr S Aleksic of Institute of Hygiene, Germany
- j. Dr J Lefebvre of Bacterial Identification Section, Laboratoire de Santé Publique du Québec, Canada
- k. Strains C600 and WG1 from Dr. B.J. Backmann of Department of Biology, Yale University, USA.

TABLE 7 PCR assay result using primers based on the *E. coli* serotype O16 (strain K-12) O antigen gene cluster sequence

Gene	Function	Base positions of the gene	Forward primer (base positions)	Reverse primer (base positions)	Length of the PCR fragment	Number of pools (out of 21) giving band of correct size	Annealing temperature of the PCR
<i>rmlB</i> *	TDP-rhamnose pathway	90-1175	#1064(91-109)	#1065(1175-1157)	1085bp	17	60°C
<i>rmlD</i> *	TDP-rhamnose pathway	1175-2074	#1066(1175-1193)	#1067 (2075-2058)	901bp	13	60°C
<i>rmlA</i> *	TDP-rhamnose pathway	2132-3013	#1068(2131-2148)	#1069(3013-2995)	883bp	0	60°C
<i>rmlC</i> *	TDP-rhamnose pathway	3013-3570	#1070(3012-3029)	#1071(3570-3551)	559bp	0	60°C
<i>gtf</i> *	Galactofuranose pathway	4822-5925	#1074(4822-4840)	#1075(5925-5908)	1104bp	0	55°C
<i>wzx</i> *	Flippase	3567-4814	#1072(3567-3586)	#1073(4814-4797)	1248bp	0	55°C
<i>wzy</i> *	O polymerase	5925-7091	#1076(5925-5944)	#1077(7091-7074)	1167bp	0	60°C
<i>wbbI</i> *	Galactofuranosyl transferase	7094-8086	#1078 (7094-7111)	#1079(8086-8069)	993bp	0	50°C
<i>wbbJ</i> *	Acetyltransferase	8067-8654	#1080(8067-8084)	#1081(8654-8632)	588bp	0	60°C
<i>wbbK</i> **	Glucosyl transferase	5770-6888	#1082(5770-5787)	#1083(6888-6871)	1119bp	0	55°C
<i>wbbL</i> ***	Rhamnosyltransferase	679-1437	#1084(679-697)	#1085(1473-1456)	795bp	0****	55°C

* ** *** Base positions based on GenBank entry U09876, U03041 and L19537 respectively
 , , ,

 19 pools giving a band of wrong size

TABLE 8 PCR assay data using 0111 primers

Gene	Base positions of the gene according to SEQ ID NO: 1	Forward primer (base positions)	Reverse primer (base positions)	Length of the PCR fragment	Number of pools (out of 21) giving band of correct size	Annealing temperature of the PCR
<i>wbdH</i>	739-1932	#866 (739-757)	#867(1941-1924)	1203bp	0	60°C
		#976(925-942)	#978(1731-1714)	807bp	0	60°C
		#976(925-942)	#979(1347-1330)	423bp	0	60°C
		#977(1165-1182)	#978(1731-1714)	567bp	0	60°C
<i>wzx</i>	8646-9911	#969(8646-8663)	#970(9908-9891)	1263bp	0	50°C
		#1060(8906-8923)	#1062(9468-9451)	563bp	0	60°C
		#1061(9150-9167)	#1063 (9754-9737)	605bp	0	50°C
<i>wzy</i>	9901-10953	#900(9976-9996)	#901(10827-10807)	852bp	0	60°C
		#980(10113-10130)	#983(10484-10467)	372bp	0*	61°C
<i>wbdL</i>	10931-11824	#870(10931-10949)	#871(11824-11796)	894bp	7	60°C
<i>wbdM</i>	11821-12945	#868(11821-11844)	#869(12945-12924)	1125bp	0	60°C
		#984(12042-12059)	#987(12447-12430)	406bp	0	60°C
		#985(12258-12275)	#986(12698-12681)	441bp	0**	65°C

* Giving a band of wrong size in all pools

** One pool giving a band of wrong size

TABLE 8A PCR specificity test data using 0111 primers

Gene	Base positions of the gene according to SEQ ID NO: 1	Forward primer (base positions)	Reverse primer (base positions)	Length of the PCR fragment	Number of pools (pools no. 25-28) giving band of correct size	Annealing temperature of the PCR
<i>wbdH</i>	739-1932	#866 (739-757)	#867(1941-1924)	1203bp	0*	60°C
		#976(925-942)	#978(1731-1714)	807bp	0	60°C
		#976(925-942)	#979(1347-1330)	423bp	0	60°C
		#977(1165-1182)	#978(1731-1714)	567bp	0	60°C
<i>wzx</i>	8646-9911	#969(8646-8663)	#970(9908-9891)	1263bp	0	55°C
		#1060(8906-8923)	#1062(9468-9451)	563bp	0	60°C
		#1061(9150-9167)	#1063 (9754-9737)	605bp	0*	50°C
<i>wzy</i>	9901-10953	#900(9976-9996)	#901(10827-10807)	852bp	0	60°C
		#980(10113-10130)	#983(10484-10467)	372bp	0**	60°C
<i>wbdL</i>	10931-11824	#870(10931-10949)	#871(11824-11796)	894bp	0	60°C
<i>wbdM</i>	11821-12945	#868(11821-11844)	#869(12945-12924)	1125bp	0	60°C
		#984(12042-12059)	#987(12447-12430)	406bp	0	60°C
		#985(12258-12275)	#986(12698-12681)	441bp	0*	65°C

* 1 pool giving a band of wrong size

** 2 pools giving 3 bands of wrong sizes, 1 pool giving 2 bands of wrong sizes

TABLE 9 PCR results using primers based on the *E. coli* O157 sequence

Gene	Function	Base position of the gene according to SEQ ID NO: 2	Forward primer (base positions)	Reverse primer (base positions)	Length of the PCR fragment	Number of pools (out of 21) giving band of correct size	Annealing temperature of the PCR
<i>wbdN</i>	Sugar transferase	79-861	#1197(79-96)	#1198 (861-844)	783	0	55°C
			#1199(184-201)	#1200(531-514)	348	0	55°C
			#1201(310-327)	#1202(768-751)	459	0	55°C
<i>wzy</i>	O antigen	858-2042	#1203(858-875)	#1204(2042-2025)	1185	0*	50°C
			#1205(1053-1070)	#1206(1619-1602)	567	0	63°C
			#1207(1278-1295)	#1208(1913-1896)	636	0	60°C
<i>wbdO</i>	Sugar transferase	2011-2757	#1209(2011-2028)	#1210(2757-2740)	747	0	50°C
			#1211(2110-2127)	#1212(2493-2476)	384	0**	62°C
			#1213(2305-2322)	#1214(2682-2665)	378	0	60°C
<i>wzx</i>	O antigen flippase	2744-4135	#1215(2744-2761)	#1216(4135-4118)	1392	0	50°C
			#1217(2942-2959)	#1218(3628-3611)	687	0***	63°C
<i>wbdP</i>	Sugar transferase	5257-6471	#1221(5257-5274)	#1222(6471-6454)	1215	0	55°C
			#1223(5440-5457)	#1224(5973-5956)	534	0	55°C
			#1225(5707-5724)	#1226(6231-6214)	525	0	55°C
<i>wbdR</i>	N-acetyl	13156-13821	#1229(13261-13278)	#1230(13629-13612)	369	0	55°C
			#1231(13384-13401)	#1232(13731-13714)	348	0	60°C

* 3 bands of wrong size in one pool, 1 band of wrong size in all other pools

** 3 bands of wrong sizes in 9 pools, 2 bands of wrong size in all other pools

*** 2 bands of wrong sizes in 2 pools, 1 band of wrong size in 7 pools

TABLE 9A PCR results using primers based on the *E. coli* O157 sequence

Gene	Function	Base position of the gene according to SEQ ID NO: 2	Forward primer (base positions)	Reverse primer (base positions)	Length of the PCR fragment	Number of pools (pools no. 37-42) giving band of correct size	Annealing temperature of the PCR
<i>wbdN</i>	Sugar transferase	79-861	#1197(79-96)	#1198 (861-844)	783	0*	55°C
			#1199(184-201)	#1200(531-514)	348	0*	55°C
			#1201(310-327)	#1202(768-751)	459	0	61°C
<i>wzy</i>	O antigen	858-2042	#1203(858-875)	#1204(2042-2025)	1185	1**	50°C
			#1205(1053-1070)	#1206(1619-1602)	567	0***	60°C
			#1207(1278-1295)	#1208(1913-1896)	636	0	60°C
<i>wbdO</i>	Sugar transferase	2011-2757	#1209(2011-2028)	#1210(2757-2740)	747	0	50°C
			#1211(2110-2127)	#1212(2493-2476)	384	0****	61°C
			#1213(2305-2322)	#1214(2682-2665)	378	0	60°C
<i>wzx</i>	O antigen flippase	2744-4135	#1215(2744-2761)	#1216(4135-4118)	1392	0	50°C
			#1217(2942-2959)	#1218(3628-3611)	687	0	63°C
<i>wbdP</i>	Sugar transferase	5257-6471	#1221(5257-5274)	#1222(6471-6454)	1215	0	55°C
			#1223(5440-5457)	#1224(5973-5956)	534	0*	60°C
			#1225(5707-5724)	#1226(6231-6214)	525	0	55°C
<i>wbdR</i>	N-acetyl transferase	13156-13821	#1229(13261-13278)	#1230(13629-13612)	369	0	50°C
			#1231(13384-13401)	#1232(13731-13714)	348	0	60°C

* 1 band of wrong size in one pool

** pool #39 giving two bands, one band of correct size, the other band of wrong size in another pool.

*** 2 bands of wrong sizes in one pool

**** 3 bands of wrong sizes in 2 pools, 2 bands of wrong sizes in 2 other pools

CLAIMS:

1. A nucleic acid molecule which encodes all or part of an *E. coli* flagellin protein, the molecule being capable of identifying the H serotype of an *E. coli* when hybridised to a gene of the *E. coli* which encodes a flagellin protein, provided that the molecule does not encode a flagellin protein expressed by the *E. coli* H1, H7, H12 or H48 type strains.
2. A nucleic acid molecule according to claim 1 wherein the molecule is derived from a *fliC* gene.
3. A nucleic acid molecule according to claim 1 including all or part of a sequence according to any one of SEQ ID NOs:1 to 68.
4. A nucleic acid molecule according to claim 1 consisting of all or part of a sequence according to any one of SEQ ID NOs: 1 to 68.
5. A nucleic acid molecule according to claim 4 wherein the molecule is from about 10 to 20 nucleotides in length.
6. A primer selected from the group of primers shown in Table 3.
7. A method of detecting the H serotype of *E. coli* in a sample, the method comprising the following steps:
- (a) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1 in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and
 - (b) detecting a nucleic acid molecule which is hybridised to the gene, to detect the H serotype of the *E. coli* in the sample.

8. A method according to claim 7 wherein the hybridised nucleic acid molecules are detected by Southern Blot analysis.

5

9. A method of detecting the H serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of an *E. coli* in the sample with a pair of nucleic acid molecules according to claim 1 in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene; and

(b) detecting a pair of nucleic acid molecules which is hybridised to the gene, to detect the H serotype of the *E. coli* in the sample.

15

10. A method according to claim 9 wherein the hybridised pairs of nucleic acid molecules are detected by the polymerase chain reaction.

20

11. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

(a) contacting a gene of the *E. coli* with a nucleic acid molecule derived from a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a *E. coli* O antigen, in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene;

25

(b) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1 in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and

30

(c) detecting nucleic acid molecules which are hybridised to the genes, to detect the H and O serotype of the *E. coli* in the sample.

35

- 85 -

12. A method according to claim 11 wherein the nucleic acid molecule of step (a) is selected from the group consisting of:

5 *wbdH* (nucleotide position 739 to 1932 of Figure 5),
 wzx (nucleotide position 8646 to 9911 of Figure 5),
 wzy (nucleotide position 9901 to 10953 of Figure 5),
 wbdM (nucleotide position 11821 to 12945 of Figure 5),
 wbdN (nucleotide position 79 to 861 of Figure 6),
 wbdO (nucleotide position 2011 to 2757 of Figure 6),
10 *wbdP* (nucleotide position 5257 to 6471 of Figure 6),
 wbdR (nucleotide position 13156 to 13821 of Figure 6),
 wzx (nucleotide position 2744 to 4135 of Figure 6) and
 wzy (nucleotide position 858 to 2042 of Figure 6).

15 13. A method according to claim 12 wherein the nucleic acid molecule of step (a) is a primer selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

20 14. A method according to claim 11 wherein the hybridised nucleic acid molecules are detected by Southern Blot analysis.

25 15. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

30 (a) contacting a gene of the *E. coli* with a pair of nucleic acid molecules derived from a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a *E. coli* O antigen, in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene;

35 (b) contacting a gene of an *E. coli* in the sample with a pair of nucleic acid molecules according to claim 1 in conditions sufficient to allow the pair of nucleic acid molecules to hybridise to the gene; and

 (c) detecting pairs of nucleic acid molecules which

- 86 -

are hybridised to the genes, to detect the H and O serotype of the *E. coli* in the sample.

5 16. A method according to claim 15 wherein the pair of nucleic acid molecules of step (a) is selected from the group consisting of:

wbdH (nucleotide position 739 to 1932 of Figure 5),
wzx (nucleotide position 8646 to 9911 of Figure 5),
wzy (nucleotide position 9901 to 10953 of Figure 5),
10 *wbdM* (nucleotide position 11821 to 12945 of Figure 5),
wbdN (nucleotide position 79 to 861 of Figure 6),
wbdO (nucleotide position 2011 to 2757 of Figure 6),
wbdP (nucleotide position 5257 to 6471 of Figure 6),
wbdR (nucleotide position 13156 to 13821 of Figure 6),
15 *wzx* (nucleotide position 2744 to 4135 of Figure 6) and
wzy (nucleotide position 858 to 2042 of Figure 6).

20 17. A method according to claim 15 wherein the nucleic acid molecules of the pair of step (a) are primers selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

25 18. A method according to claim 15 wherein the hybridised pairs of nucleic acid molecules are detected by the polymerase chain reaction.

19. A method for detecting the H and O serotype of *E. coli* in a sample, the method comprising the following steps:

30 (a) contacting a gene of an *E. coli* in the sample with a nucleic acid molecule according to claim 1, in conditions sufficient to allow the nucleic acid molecule to hybridise to the gene; and

35 (b) detecting a nucleic acid molecule which is hybridised to the gene, to detect the H and O serotype of *E. coli* in the sample.

- 87 -

20. A method according to claim 19 wherein the nucleic acid molecule is according to any one of SEQ ID NOS: 9, 55, 57 to 65.

5 21. A method according to any one of claims 8, 9, 11, 15 or 19 wherein the sample is selected from the group consisting of a sample derived from food, a sample derived from faeces and a sample derived from a patient or animal.

10 22. A kit for identifying the H serotype of *E. coli*, the kit comprising at least one nucleic acid molecule according to any one of claims 1 to 6.

15 23. A kit for identifying the H and O serotype of *E. coli*, the kit comprising:

(a) at least one nucleic acid molecule according to any one of claims 1 to 6; and

20 (b) at least one nucleic acid molecule derived from and specific for a gene encoding a transferase or a gene encoding an enzyme for the transport or processing of a polysaccharide or oligosaccharide unit, the gene being involved in the synthesis of a particular *E. coli* O antigen.

25 24. A kit according to claim 23 wherein the at least one nucleic acid molecule of (a) is selected from the group consisting of:

30 *wbdH* (nucleotide position 739 to 1932 of Figure 5),
wzx (nucleotide position 8646 to 9911 of Figure 5),
wzy (nucleotide position 9901 to 10953 of Figure 5),
wbdM (nucleotide position 11821 to 12945 of Figure 5),
wbdN (nucleotide position 79 to 861 of Figure 6),
wbdO (nucleotide position 2011 to 2757 of Figure 6),
wbdP (nucleotide position 5257 to 6471 of Figure 6),
35 *wbdR* (nucleotide position 13156 to 13821 of Figure 6),
wzx (nucleotide position 2744 to 4135 of Figure 6) and
wzy (nucleotide position 858 to 2042 of Figure 6).

- 88 -

25. A kit according to claim 24 wherein the nucleic acid molecule of (a) is a primer selected from the group of primers shown in Tables 8, 8A, 9 and 9A.

5

AMENDED SHEET
IPER/100

1/96

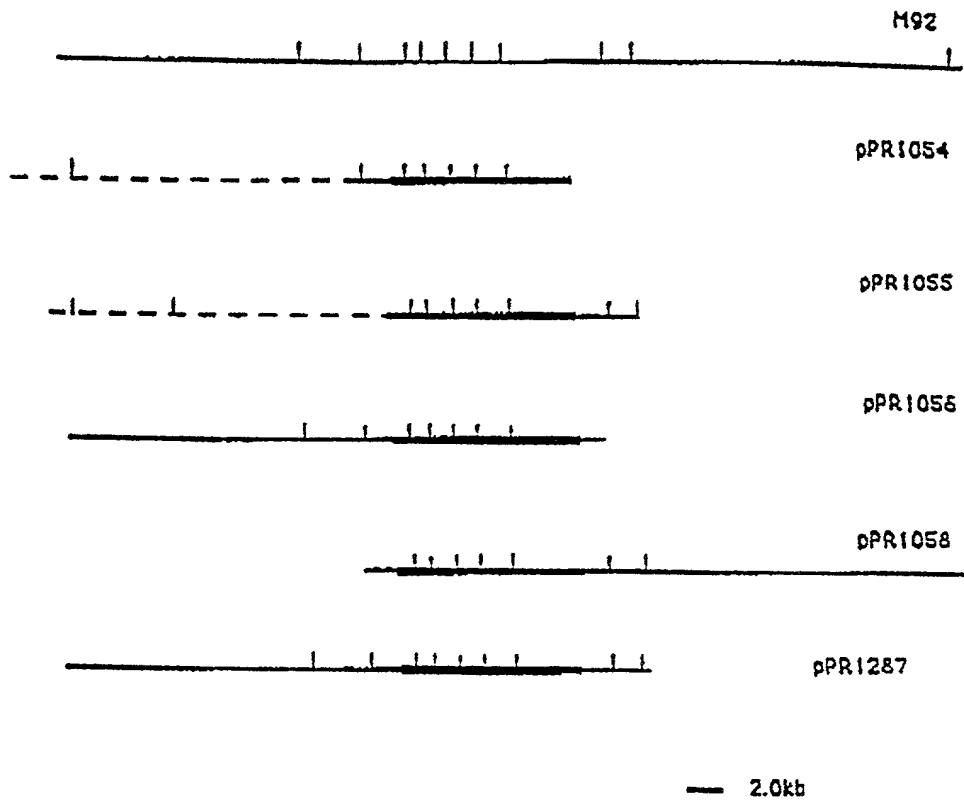


Figure 1

2/96

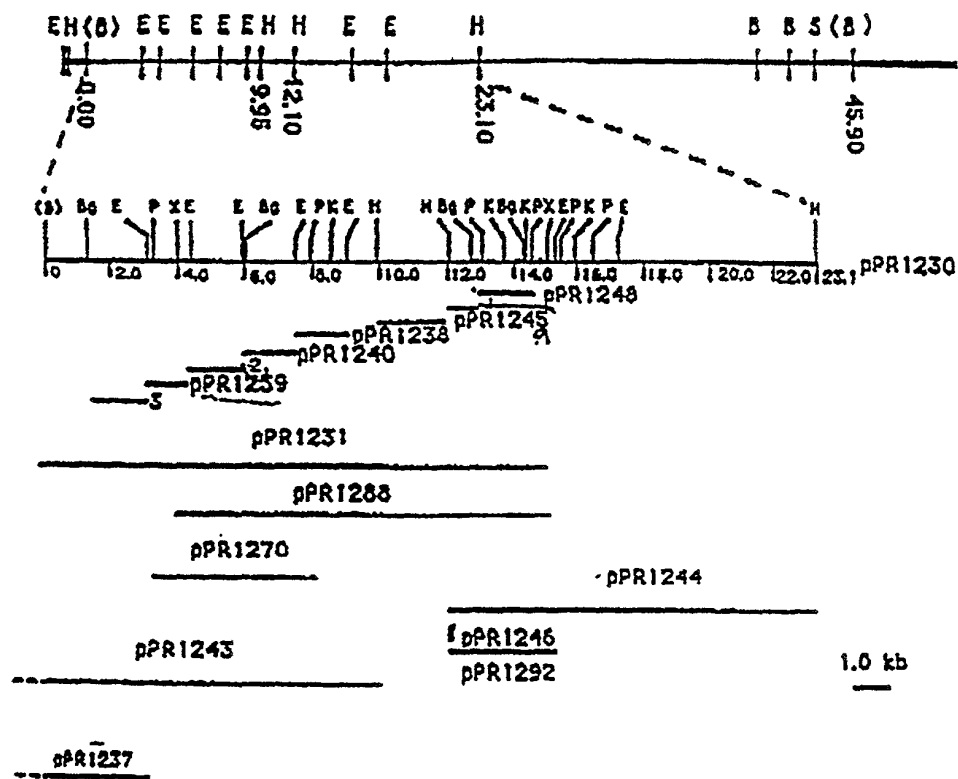


Figure 2

WO 99/61458

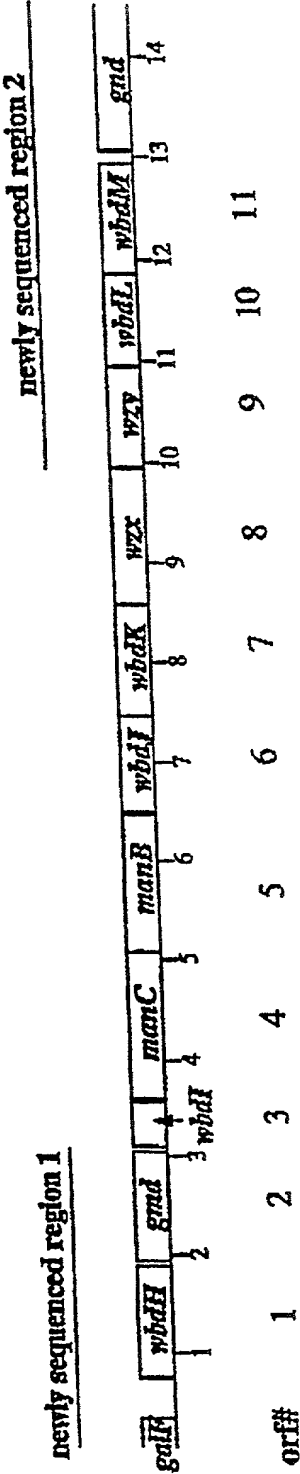


Figure 3

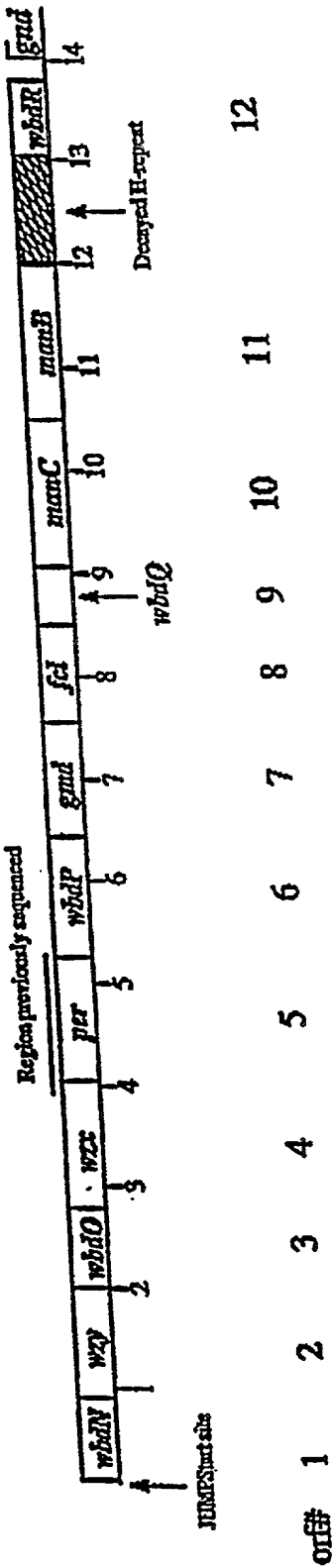


Figure 4

5/96

GATCTGATGGCCGTAGGGCGCTACGTGCTTCTGCTGATATCTGGGCTGAGTTGGAAAAA 60
 ACTGCTCCAGGTGCCTGGGGACGTATTCAACTGACTGATGCTATTGCAGAGTTGGCTAAA 120
 AAACAGTCTGTTGATGCCATGCTGATGACCGGCGACAGCTACGACTGCGGTAAGAAGATG 180
 GGCTATATGCAGGCATTCGTTAAGTATGGGCTGCGCAACCTTAAAGAAGGGGCGAAGTTC 240
 CGTAAGAGCATCAAGAAGCTACTGAGTGAGTAGAGATTACACGTCTTTGTGACGATAAG 300
 CCAGAAAAAATAGCGGCAGTTAACATCCAGGCTTCTATGCTTTAAGCAATGGAATGTTAC 360
 TGCCGTTTTTTTATGAAAAATGACCAATAATAACAAGTTAACCTACCAAGTTTAATCTGCT 420
 TTTTGTGTGGATTTTTTCTTGTCTGCTGGTCGCATTGGTAAGACAATTAGCGTGAGTTTTA 480
 GAGAGTTTTGCGGGATCTCGCGGAAGTGTCTCACATCTTTGGCATTAGTTAGTGCACCTGG 540
 TAGCTGTTAAGCCAGGGGCGGTAGCTTGCCTAATTAATTTTAAAGTATACATTTATTCT 600
 TGCCGCTTATAGCAAATAAAGTCAATCGGATTAACTTCTTTTCCATTAGGTAAAAGAGT 660
 GTTTGTAGTCGCTCAGGGAAATTGGTTTTGGTAGTAGTACTTTTCAAATTATCCATTTTC 720

Start of orf1

M L L C C I H I N V Y Y L L
 CGATTTAGATGGCAGTTGATGTTACTATGCTGCATACATATCAATGTATATTATTTACTT 780
 L E C D M K K I V I I G N V A S M M L R
 TTAGAATGTGATATGAAAAAATAGTGATCATAGGCAATGTAGCGTCAATGATGTTAAGG 840
 F R K E L I M N L V R Q G D N V Y C L A
 TTCAGGAAAGAATTAATCATGAATTTAGTGAGGCAAGGTGATAATGTATATTGTCTAGCA 900
 N D F S T E D L K V L S S W G V K G V K
 AATGATTTTTTCCACTGAAGATCTTAAAGTACTTTCTGTCATGGGGCGTTAAGGGGGTTAA 960
 F S L N S K G I N P F K D I I A V Y E L
 TTCTCTCTTAACTCAAAGGTATTAATCCTTTTAAAGGATATAATTGCTGTTTATGAAC 1020
 K K I L K D I S P D I V F S Y F V K P V
 AAAAAAATCTTAAGGATATTTCCCAGATATTGTATTTTCATATTTTGTAAGCCAGTA 1080
 I F G T I A S K L S K V P R I V G M I E
 ATATTTGGAATATTGCTTCAAAGTTGTCAAAGTGCCAAGGATTGTTGGAATGATTGAA 1140
 G L G N A F T Y Y K G K Q T T K T K M I
 GGTCTAGGTAATGCCTTCACTTATTATAAGGGAAAGCAGACCACAAAACTAAAATGATA 1200
 K W I Q I L L Y K L A L P M L D D L I L
 AAGTGGATACAAATCTTTTATATAAGTTAGCATTACCGATGCTTGATGATTTGATTCTA 1260
 L N H D D K K D L I D Q Y N I K A K V T
 TTAAATCATGATGATAAAAAAGATTTAATCGATCAGTATAATATTAAGCTAAGGTAACA 1320
 V L G G I G L D L N E F S Y K E P P K E
 GTGTTAGGTGGGATTGGATTGGATCTTAATGAGTTTTCATATAAAGAGCCACCGAAAGAG 1380
 K I T F I F I A R L L R E K G I F E F I
 AAAATTACCTTTATTTTATAGCAAGGTTATTAAGAGAGAAAGGGATATTTGAGTTTATT 1440
 E A A K F V K T T Y P S S E F V I L G G
 GAAGCCGCAAAGTTCGTTAAGACAACCTATCCAAGTTCTGAATTTGTAATTTTAGGAGGT 1500

6/96

F E S N N P F S L Q K N E I E S L R K E 1560
 TTTGAGAGTAATAATCCTTTCTCATTACAAAAAATGAAATTGAATCGCTAAGAAAAGAA
 H D L I Y P G H V E N V Q D W L E K S S 1620
 CATGATCTTATTTATCCTGGTCATGTGGAAAAATGTTCAAGATTGGTTAGAGAAAAGTTCT
 V F V L P T S Y R E G V P R V I Q E A M 1680
 GTTTTGTGTTTACCTACATCATATCGAGAAGGCGTACCAAGGGTGATCCAAGAAGCTATG
 A I G R P V I T T N V P G C R D I I N D 1740
 GCTATTGGTAGACCTGTAATAACAATAATGTACCTGGGTGTAGGGATATAATAAATGAT
 G V N G F L I P P F E I N L L A E K M K 1800
 GGGGTCAATGGCTTTTGTATACCTCCATTTGAAATTAATTTACTGCGAGAAAAAATGAAA
 Y F I E N K D K V L E M G L A G R K F A 1860
 TATTTTATTGAGAATAAAGATAAAGTACTCGAAATGGGGCTTGCTGGAAGGAAGTTGCA
 E K N F D A F E K N N R L A S I I K S N 1920
 GAAAAAACCTTTGATGCTTTTGAATAAATAATAGACTAGCATCAATAATAAATCAAAT

 N D F * **End of orf1**
 AATGATTTTTGACTTGAGCAGAAATTATTTATATTTCAATCTGAAAAATAAAGGCTGTTA 1980

Start of orf2
 M N K V A L I T G I T G Q D G S Y L A 2040
 TTATGAATAAAGTGGCATTAATTACTGGTATCACTGGGCAAGATGGCTCCTATTTGGCAG
 E L L L E K G Y E V H G I K R R A S S F 2100
 AATTATTGTTAGAAAAAGGTTATGAAGTTCATGGTATTAAACGCCGTGCATCTTCATTTA
 N T E R V D H I Y Q D S H L A N P K L F 2160
 ATACTGAGCGAGTGGATCACATCTATCAGGATTCACATTTAGCTAATCCTAACTTTTTTC
 L H Y G D L T D T S N L T R I L K E V Q 2220
 TACACTATGGCGATTTGACAGATACTTCCAATCTGACCCGTATTTTAAAAGAAGTTCAAC
 P D E V Y N L G A M S H V A V S F E S P 2280
 CAGATGAAGTTTACAATTTGGGGGCGATGAGCCATGTAGCGGTATCATTTGAGTCAACAG
 E Y T A D V D A I G T L R L L E A I R I 2340
 AATACACTGCTGATGTTGATGCGATAGGAACATTGCGTCTTCTTGAAGCTATCAGGATAT
 L G L E K K T K F Y Q A S T S E L Y G L 2400
 TGGGGCTGGAAAAAAGACAAAATTTTATCAGGCTTCAACTTCAGAGCTTTATGGTTTGG
 V Q E I P Q K E T T P F Y P R S P Y A V 2460
 TTCAAGAAATTCCACAAAAGAGACTACGCCATTTTATCCACGTTGCGCTTATGCTGTTG
 A K L Y A Y W I T V N Y R E S Y G M F A 2520
 CAAAATTATATGCCTATTGGATCACTGTTAATTATCGTGAGTCTTATGGTATGTTTGCCT
 C N G I L F N H E S P R R G E T F V T R 2580
 GCAATGGTATTCTCTTTAACCACGAATCACCTCGCCGTGGCGAGACCTTTGTTACTCGTA
 K I T R G I A N I A Q G L D K C L Y L G 2640
 AAATAACACGCGGGATAGCAAATATTGCTCAAGGTCTTGATAAATGCTTATACTTGGGAA
 N M D S L R D W G H A K D Y V K M Q W M 2700
 ATATGGATTCTCTGCGTGATTGGGGACATGCTAAGGATTATGTCAAAATGCAATGGATGA

7/96

M L Q Q E T P E D F V I A T G I Q Y S V
 TGCTGCAGCAAGAACTCCAGAAGATTTTGTAAATGCTACAGGAATTCAATATTCTGTCC 2760
 R E F V T M A A E Q V G I E L A F E G E
 GTGAGTTTGTACAAATGGCGGCAGAGCAAGTAGGCATAGAGTTAGCATTGGAAGGTGAGG 2820
 G V N E K G V V V S V N G T D A K A V N
 GAGTAAATGAAAAAGGTGTTGTTGTTTCGGTCAATGGCACTGATGCTAAAGCTGTAAACC 2880
 P G D V I I S V D P R Y F R P A E V E T
 CGGGCGATGTAATTATATCTGTAGATCCAAGGTATTTTAGGCCTGCAGAAGTTGAAACCT 2940
 L L G D P T N A H K K L G W S P E I T L
 TGCTTGGCGATCCTACTAATGCGCATAAAAAATTAGGATGGAGCCCTGAAATTACATTGC 3000
 R E M V K E M V S S D L A I A K K N V L
 GTGAAATGGTAAAAGAAATGCTTTCAGCGGATTTAGCAATAGCGAAAAGAAAGCTCTTGC 3060

End of orf2

L K A N N I A T N I P Q E *
 TGAAGCTAATAACATTGCCACTAATATTCGCAAGAA TAAAAAGATAATACATTAAAT 3120

Start of orf3

M F

AATTAAAAATGGTGCTAGATTTATTAGTACCATTATTTTTTTTTTGGGTGACTAATGTTTA 3180
 I T S D K F R E I I K L V P L V S I D L
 TTACATCAGATAAAATTTAGAGAAATTTATCAAGTTAGTTCCATTAGTATCAATTGATCTGC 3240
 L I E N E N G E Y L F G L R N N R P A K
 TAATTGAAACGAGAAATGGTGAATATTTATTTGGTCTTAGGAATAATCCAGCGGCAAAA 3300
 N Y F F V P G G R I R K N E S I K N A F
 ATTAATTTTTTGTTCAGGCTGGTAGGATTCCGCAAAAATGAATCTATTAAAAATGCTTTTA 3360
 K R I S S M E L G K E Y G I S G S V F N
 AAAGATATCATCTATGGAATTAGGTAAGAGTATGCTATTTTCAGGAAGTCTTTTTTAATG 3420
 G V W E H F Y D D G F F S E G E A T H Y
 GTGTATGGCAACATTTCTATGATGATGTTTTTTTTTCTGAAGGCGAGGCAACACATTATA 3480
 I V L C Y T L K V L K S E L N L P D D Q
 TAGTGGTTTGTACACACTGAAAGTTCTTAAAGTGAATTCGAATCTCCAGATGATCAAC 3540
 H R E Y L W L T K H Q I N A K Q D V H N
 ATCGTGAATACCTTTGGCTAACTAAACACCAATAAATGCTAAACAAGATCTTCATAACT 3600

End of orf3

Start of orf4

M

Y S K N Y F L *
 ATTCAAAAATTAATTTTTTGTAAATTTTATTAAAAATTAATATGCGAGAGAAATGTAATG 3660
 S Q C L Y P V I I A G G T G S R L W P L
 CTCATGCTCTTTACCTGTAAATATTGCGCGAGGAACCGGAAGCGGTCTATGCGCGTTGT 3720
 S R V L Y P K Q F L N L V G D S T M L Q
 CTCGAGTATTATACCTAAACAATTTTTAAATTTAGTTGGGGATTCTACAAATGTTGCAAA 3780
 T T I T R L D G I E C E N P I V I C N E
 CAACAATTACGCGTTTGCATGCCATGGAATGCGAAATCCAAATGTTATCTGCAATGAAG 3840
 D H R F I V A E Q L R Q I G K L T K N I
 ATCACCGATTTATTGTAGCAGAGCAATTACGACAGATTGCTAAGCTAAACCAAGAAATATTA 3900
 I L E P K G R N T A P A I A L A A F I A
 TAATTGAGCGGAAGGCGTAAATACTGCACTGCGATAGCTTTAGCTGCTTTTATCGCTC 3960

Q K N N P N D D P L L L V L A A D H S I 4020
 AGAAGAATAATGCTAATGACGACCCCTTTATTATTAGTACTTGCGGCAGACCACTCTATAA
 N N E K A F R E S I I K A M P Y A T S G 4080
 ATAATGAAAAAGCATTTCGAGAGTCAATAATAAAAGCTATGCCGTATGCAACTTCTGCGA
 K L V T F G I I P D T A N T G Y G Y I K 4140
 AGTTAGTAACATTTCGAATTATTCCGGACACGGCAATACTGGTTATGGATATATTAAGA
 R S S S A D P N K E F P A Y N V A E F V 4200
 GAAGTTCTTCAGCTGATCCTAATAAGAATTCCCGAGCATATAATGTTGCGGAGTTTGTAG
 E K P D V K T A Q E Y I S S G N Y Y W N 4260
 AAAAACCAGATGTTAAACAGCAGCAATATATTTCCAGTGGGAATTATTACTGGAATA
 S G M F L F R A S K Y L D E L R K F R P 4320
 GCGAATGTTTTTATTTCCGCCAGTAAATATCTTGATGAACCTACGGAAATTTAGACCAG
 D I Y H S C E C A T A T A N I D M D F V 4380
 ATATTTATCATAGCTGTGAATGTGCAACCCCTACAGCAATATAGATATGGACTTTGTCC
 R I N E A E F I N C P E E S I D Y A V M 4440
 GAATTAACGAGGCTGAGTTTATTAATTGTCTCTGAAGAGTCTATCGATTATGCTGTGATGG
 E K T K D A V V L P I D I G W N D V G S 4500
 AAAAAACAAAGACGCTGTAGTTCTTCCGATAGATATTGCTGGAATGACGTGGGTTCTT
 W S S L W D I S Q K D C H G N V C H G D 4560
 GGTCATCACTTTTGGGATATAAGCCAAAAGGATTGCCATGGTAATGTGTGCCATGGGGATG
 V L N H D G E N S F I Y S E S S L V A T 4620
 TGCTCAATCATGATGGAGAAAATAGTTTATTTACTCTGAGTCAAGTCTGGTTGGGACAG
 V G V S N L V I V Q T K D A V L V A D R 4680
 TCGGAGTAAGTAATTTAGTAATTTGTCCAAACCAAGGATGCTGTACTGGTTGCGGACCCGTG
 D K V Q N V K N I V D D L K K R K R A E 4740
 ATAAAGTCCAAAATGTTAAAAACATAGTTGACGATCTAAAAAGAGAAAACGTGCTGAAT
 Y Y M H R A V F R P W G K F D A I D Q G 4800
 ACTACATGCATCGTGCAGTTTTTCCGCCCTTGGGGTAAATTCGATGCAATAGACCAAGGCG
 D R Y R V K K I I V K P G E G L D L R M 4860
 ATAGATATAGAGTAATAAAAAATAATAGTTAAACCAGGAGAGGGTTAGATTTAAGGATGC
 H H H R A E H W I V V S G T A K V S L G 4920
 ATCATCATAGGGCAGAGCATTGGATTGTTGTATCCCGTACTGCTAAAGTTTCACTAGCTA
 S E V K L L V S N E S I Y I P Q G A K Y 4980
 GTGAAGTTAACTATTAGTTTCTAATGAGTCTATATATATATCCCTCAGGGAGCAAAATATA
 S L E N P G V I P L H L I E V S S G D Y 5040
 GTCTTGAGAAATCCAGGCGTAATACCTTTGCATCTAATTGAAGTAAGTTCTGGTGATTACC
 L E S D D I V R F T D R Y N S K Q F L K 5100
 TTGAATCAGATGATATAGTGGCTTTTACTGACAGATATAACAGTAACAATTCCTAAAGC

End of orf4 Start of orf5

M N K I T C F K A Y D I R G R L

R D *

GAGATTGATAAATAATGAATAAAATAAAGCTTTCAGCTTCAAGCATATGATATAAGTGGGCGCTCT—5160

9/96

G A E L N D E I A Y R I G R A Y G E F F 5220
 TGCTGCTGAATTGAATGATGAAATAGCATATAGAATTGGTCGGCCTTATGGTGAGTTTTT
 K P Q T V V V G G D A R L T S E S L K K 5280
 TAAACCTCAAACTGTAGTTGTGCGGAGGAGATGCTGCTTAAACAGTGAGAGTTTAAAGAA
 S L S N G L C D A G V N V L D L G M C G 5340
 ATCACTCTCAAAATGGGCTATGTGATGCGAGGCGTAAATGTCTTAGATCTTGGAAATGTGTGG
 T E E I Y F S T W Y L G I D G G I E V T 5400
 TACTGAAGAGATATATTTTTCCACTTGGTATTTAGGAATTGATGCTGGAATCGAGGTAAAC
 A S H N P I D Y N G M K L V T K G A R P 5460
 TGCAAGCCATAATCCAATTGATTATAATGGAATGAAATTAGTAACCAAGCGTGTCTGAGCC
 I S S D T G L K D I Q Q L V E S N N F E 5520
 AATCAGCAGTGACACAGGTCTCAAGATATACAACAATTAGTAGAGAGTAATAATTTTTGA
 E L N L E K K G N I T K Y S T R D A Y I 5580
 AGAGCTCAACCTAGAAAAAAGGGAATATTACCAAAATATCCACCCGAGATGCCTACAT
 N H L M G Y A N L Q K I K K I K I V V N 5640
 AATCATTTTGATGGGCTATGCTAATCTGCAAAAAATAAAAAAATCAAAATAGTTGTGAA
 S G N G A A G P V I D A I E E C F L R N 5700
 TTCTGGGAATGGTGCAGCTGGTCTCTTATTGATGCTATTGAGGAATGCTTTTTTACGGAA
 N I P I Q F V K I N N T P D G N F P H G 5760
 CAATATTCGGATTTCAGTTTGTAAAAATAAATAACACCCGATGCTAATTTTCCACATGG
 I P N P L L P E C R E D T S S A V I R H 5820
 TATCCCTAATCCATTACTACCTGAGTGCAGAGAGATACCAGCAGTGGGTTATAAGACA
 S A D F G I A F D G D F D R C F F F D E 5880
 TAGTGGCTGATTTTGCTATTGCAATTTGATGGTGATTTTGATAGCTGTTTTTCTTTTGATGA
 N G Q F I E G Y Y I V G L L A E V F L G 5940
 AATGCGACAATTTATTGAAGGATACTACATTGTTGGTTATTAGCGGAAGTTTPTTTAGG
 K Y P N A K I I H D P R L I W N T I D I 6000
 GAAATATCCAAACGCAAAATCATTCATGATCCTCGCCTTATATGCAATACTATTGATAT
 V E S H G G I P I M T K T G H A Y I K Q 6060
 CCTAGAAAGTCTGCTGGTATACCTATAATGACTAAACCGGTCATGCTTACATTAAACA
 R M R E E D A V Y G G E M S A H H Y F K 6120
 AAGAATGCGTGAAGAGGATGCGGTATATGGCGGCGAAATGAGTGGCCATCATTATTTTA
 D F A Y C D S G M I P W I L I C E L L S 6180
 AGATTTTGCATACTGCGATAGTGGAAATGATTCCTTGGATTTTAAATTTGTGAACTTTGTAG
 L T N K K L G E L V C G C I N D W P A S 6240
 TCTGACAAATAAAAAATTAGGTGAACCTGTTTGTGGTTGTATAAACGACTGGCCGGCAAG
 G E I N C T L D N P Q N E I D K L F N R 6300
 TGGAGAAATAAACTGTAGACTAGACAATCCGCAAAATGAATAGATAAAATTATTTTAAATCG
 Y K D S A L A V D Y T D G L T M E F S D 6360
 TTACAAAGATAGTGGCTTAGCTGTTGATTACACTGATGGATTAACTATGGAGTTCTCTGA
 W R F N V R C S N T E P V V R L N V E S 6420
 TTGGCGTTTTTAATGTTAGATGCTCAAAATACAGAACTGTAGTACGATTGAATGTAGAAATC
 R N N A I L M Q E K T E E I L N F I S K 6480
 TAGGAATAATGCTATTCTTATGCGAGGAAAAACAGAAAGAAATCTGAATTTTATATCAAA

11/96

L L M I A A L F F T N K P K L K R G D E
 CTGTTAATGATTGCTGCECCTTTTCTTCACTAATAAAGCAAACTTAAAGAGGTGATGAA 7680
 I I V P A V S W S T T Y Y P L Q Q Y G L
 ATAATAGTACCTGCAGTGTGTCATGGTCTACGACATATTACCCTCTGCAACAGTATGGCTTA 7740
 K V K F V D I N K E T L N I D I D S L K
 AAGCTGAAGTTTGTGCGATATCAATAAAGAACTTTAAATATTGATATCGATAGTTTGAAA 7800
 N A I S D K T K A I L T V N L L G N P N
 AATGCTATTTTCAGATAAAACAAAGCAATATTGACAGTAAATTTATTAGCTAATCCTAAT 7860
 D F A K I N E I I N N R D I I L L E D N
 GATTTTGCAAAAATAAATGAGATAATAAATAATAGGGATATTATCTTACTAGAGATAAC 7920
 C E S M G A V F Q N K Q A G T F G V M G
 TGTGAGTCCATGGGCGCGCTCTTTTCAAAATAAGCAGGACGGCACATTGCGAGTTATGGCT 7980
 T F S S F Y S H H I A T M E G G C V V T
 ACCTTTTAGTCTCTTTTACTCTCATCATATAGCTACAAATGGAAGGGGGCTGCGTAGTTACT 8040
 D D E E L Y H V L L C L R A H G W T R N
 GATGATGAAGAGCTGTATCATGTATTGTTGTGCTTTCGAGCTCATGCTTGCACAAAGAAAT 8100
 L P K E N M V T G T K S D D I F E E S F
 TTACCAAAAGAGAAATATGCTTACAGGCACTAAGAGTGTATGATATTTTTCGAAGAGTCTGTT 8160
 K F V L P G Y N V R P L E M S G A I G I
 AAGTTTCTTTTACCAGGATACAATGTTTGGCCCACTTGAATGAGTGGTGTCTATTGGGATA 8220
 E Q L K K L P G F I S T R R S N A Q Y F
 GAGCAACTTAAAAAGTTACCAGGTTTTATATCCACCAGACCTTCCAATGCACAAATTTTT 8280
 V D K F K D H P F L D I Q K E V G E S S
 GTAGATAAATTTAAAGATCATCCATTCCTTGATATACAAAAAGAGTTGGTCAAAAGTAGC 8340
 W F G F S F V I K E G A A I E R K S L V
 TGGTTTGGTTTTTCTCTCTGTATAAAGGAGGGAGCTGCTATTGACAGCAAGAGTTTAGTA 8400
 N N L I S A G I E C R P I V T G N F L K
 AATAATCTGATCTCAGCAGGCATTCGAATGCGGACCAATTGTTACTGGGAATTTTTCTCAA 8460
 N E R V L S Y F D Y S V H D T V A N A E
 AATGAACGTGTTTTGAGTTATTTTGATTACTCTGTACATGATACGGTAGCAATGCGGAA 8520
 Y I D K N G F F V G N H Q I P L F N E I
 TATATAGATAAGATCGTTTTTTTGTGCGAAACCACCAGATACCTTTCTTTAATGAATA 8580

End of orf7

D Y L R K V L K *

GATTATCTACGAAAAGTATTAAAAAACTAACCAGGGCACTCTATTTCGAATAGACTGCECT 8640

Start of orf8

M V L T V K K I L A F G Y S K V L P
 TTAAGATGCTATTAAACAGTGAATAAATTTTAGCGTTTGGCTATTCTAAAGTACTACCAC 8700

P V I E Q F V N P I C I F I I T P L I L
 CGGTTATTGAACAGTTTGTCAATGCCAATTTGCATCTTCATTATGACACCACTAATACTCA 8760

N H L G K Q S Y G N W I L L I T I V S F
 ACCACCTGGGTAAGCAAGCTATGGTAATTGGATTTTATTAATTACTATTGTATCTTTTT 8820

WO 99/61458

PCT/AU99/00385

12/96

S Q L I C G G C S A W I A K I I A E Q R 8880
 CTCAGTTAATATGTGGAGGATGTTCCGCATGCGATTGCAAAAAATCATTGCGAGAACAGAGAA
 I L S D L S K K N A L R Q I S Y N F S I 8940
 TTCTTAGTGAATTTATCAAAAAAATGCTTTACGTCAAAATTCCTATAATTTTTCAAATTG
 V I I A F A V L I S F L I L S I C F F D 9000
 TTATTATCGGCATTTGCGGTATTGATTTCTTTTCTTATATTAAGTATTGTTTCTTCGATG
 V A R N N S S F L F A I I I C G F F Q E 9060
 TTGCGAGGAATAATTCTTCATTCTTATTCCGCGATTATTATTGTTGGTTTTTTTCAGGAAG
 V D N L F S G A L K G F E K F N V S C F 9120
 TTGATAATTTATTTTAGTGGTGGCTTAAAGCTTTTGAAAAATTTAATGTATCATGTTTTT
 F E V I T R V L W A S I V I Y G I Y G N 9180
 TTGAAGTAATTACAAGAGTGCCTCTGGGCTTCTATAGTAATATATGGCATTTCACGGAATG
 A L L Y F T C L A F T I K G M L K Y I L 9240
 CACTCTTATATTTTACATGTTTACGCTTTACCATTAAGGTATGCTAAATATATTTCTTG
 V C L N I T G C F I N P N F N R V G I V 9300
 TATGTCGAAATATTACCGCTTGTTCATCAATCCTAATTTTAATAGAGTTGGGATTGTTA
 N L L N E S K W M F L Q L T G G V S L S 9360
 ATTTGTTAATGAGTCAAAATGGATGTTCTTCAATTAAGTGGTGGGCTCTCACTTAGTT
 L F D R L V I P L I L S V S K L A S Y V 9420
 TGTGATAGGCTCGTAATACCATTGATTTTATCTGTCAGTAAACTGGCTTCTTATGTCG
 P C L Q L A Q L M F T L S A S A N Q I L 9480
 CTTGCGCTCAACTAGCTCAATGATGTTCACTCTTTCTGCGTCTGCAAAATCAATATAG
 L P M F A R M K A S N T F P S N C F F K 9540
 TACCAATGTTTGTAGAAATGAAGCATCTAACACATTTCCCTCTAATTTGTTTTTTTAAA
 I L L V S L I S V L P C L A L F F F G R 9600
 TTCTGCTTGTATCACTAATTTCTGTTTTGCGCTTCTTTCGCTTATTCTTTTTTTGGTCTG
 D I L S I W I N P T F A T E N Y K L M Q 9660
 ATATATTATCAATATGGATAAACCCCTACATTTGCCAAGTGAATTTATAAATTAATGCAA
 I L A I S Y I L L S M M T S F H F L L L 9720
 TTTTAGCTATAAGTTACATTTTATTGTCATGATGACATCTTTTCATTTCTTCTTATTAG
 G I G K S K L V A N L N L V A G L A L A 9780
 GAATTGCTAAATCTAAGCTTGTGCAAAATTTAAATCTGGTTGCAAGGCTCGCACTTCTG
 A S T L I A A H Y G L Y A I S M V K I I 9840
 CTTCAACGTTAATGCGAGCTCAATTATGGCCTTTATGCAATATCTATGCTAAAAATAATAT
 Y P A F Q F Y Y L Y V A F V Y F N R A K 9900
 ATCCGGCTTTTCAATTTTATTACCTTTATGTAGCTTTTGTCTATTTTAAATAGACCGAAA

Start of orf9, End of orf8

M S I D L L F S I T E I A I V F S C T I
 N V Y *
 ATGCTATTTGATTTACTTTTTCAATTAAGTGAATCGCAATGTTTTTCTTGCCTATT 9960
 Y I F T Q C L L M R R I Y L D K S I L I 10020
 TACATATTTTACTCAATGTTGTTAATGCGGAGGATCTATTTAGATAAAAGTATTTTAATT
 L L C L L F F L V I I Q L P E L N V N G 10080
 CTTTATGCTTGCTCTTTTTTTTAGTAATCAATCACTTCCTGAGCTTAATGTAAACGGT

13/96

L V D S L K L S L P L L M V F I A F Q K 10140
 TTGGTCGATTCTTTAAAGTTATCACTGCCTTTATTGATGGTCTTTATCGCTTTTCAAAA
 P K L C L W V I I A L L F L N S A F N F 10200
 CCGAAATTATGCTTGTGGGTATTATTGCATTGTTGTTTTGAACTCTGCATTTAATTTT
 L Y L K T F D K F S S F P F T F F I L L 10260
 TTATATTTAAAGACATTCGATAAGTTTAGCTCATTTCTTTTACTTTTTTATATTGCTG
 F Y L F R L G I G N L P V Y K N K K F Y 10320
 TTTTACTTGTTTAGATTGGGAATTGGTAATTTACCGGTTTATAAAAAATAAAAAATTTTAC
 A L I F L F I L I D I M Q S L L I N Y R 10380
 GCGTTGATTTTTCTCTTTATATTAATAGACATAATGCAGTCATTGTTAATAAATTATAGG
 G Q I L Y S V I C I L I L V F K V N L R 10440
 GGCAGATTTTATATTCCGTAATTTGCATCCTGATACTTGTGTTTAAAGTTAATTTAAGA
 K K I P Y F F L M L P V L Y V I I M A Y 10500
 AAAAAGATTCCATACTTTTTTTTTAATGCTGCCAGTTTTATATGTAATTATTATGGCTTAT
 I G F N Y F N K G V T F F E P T A S N I 10560
 ATTGGTTTTAATTATTTCAATAAAGGCGTAACCTTTTTTGAACCTACAGCAAGTAATATT
 E R T G M I Y Y L V S Q L G D Y I F H G 10620
 GAACGTACGGGGATGATATATTATTGTTTCACAGCTTGGTGATTATATATTCCATGGT
 M G T L N F L N N G G Q Y K T L Y G L P 10680
 ATGGGGACATTAAATTTCTTAAATAACGGCGGACAATATAAGACGTTATATGGACTTCCA
 S L I P N D P H D F L L R F F I S I G V 10740
 TCATTAATTCCTAATGACCCTCATGATTTTTTTATTACGGTCTTTTATAAGTATTGGTGTG
 I G A L V Y H S I F F V F F R R I S F L 10800
 ATAGGAGCATTGGTTTATCATTCTATATTTTTTGTGTTTTTTTAGGAGAATATCTTTCTTA
 L Y E R N A P F I V V S C L L L L Q V V 10860
 TTATATGAGAGAAATGCTCCTTTTCATTGTTGTAAGTTGTTTGTACTGTTACAAGTTGTG
 L I Y T L N P F D A F N R L I C G L T V 10920
 TTAATTTATACATTAAACCCTTTTGATGCTTTTAATCGATTGATTGCGGGCTTACAGTT

Start of orf10

End of orf9

G V V Y G F A K I R *
 M D L Q K L D K Y T C N G N L D A
 GGAGTTGTTTATGGATTGCAAAAATTAGATAAGTATACCTGTAATGGAAATTTAGACGC 10980
 P L V S I I I A T Y N S E L D I A K C L 11040
 TCCACTTGTTTCAATAATCATTGCAACTTATAATTCTGAACTTGATATAGCTAAGTGT
 Q S V T N Q S Y K N I E I I I M D G G S 11100
 GCAATCGGTAACATAATCTTATAAGAATATTGAAATCATAAATGGATGGAGGATC
 S D K T L D I A K S F K D D R I K I V S 11160
 TTCTGATAAAACGCTTGATATTGCAAAATCGTTTAAAGACGACCGAATAAAATAGTTTC
 E K D R G I Y D A W N K A V D L S I G D 11220
 AGAGAAAGATCGTGGAATTTATGATGCCTGGAATAAAGCAGTTGATTATCCATTGGTGA
 W V A F I G S D D V Y Y H T D A I A S L 11280
 TTGGGTAGCATTATTTGGTTCAGATGATGTTTACTATCATACAGATGCAATTGCTTCATT
 M K G V M V S N G A P V V Y G R T A H E 11340
 GATGAAGGGGGTTATGGTATCTAATGGCGCCCCTGTGGTTTATGGGAGGACAGCGCACGA

14/96

G P D R N I S G F S G S E W Y N L T G F 11400
 AGGTCCCGATAGGAACATATCTGGATTTTCAGGCAGTGAATGGTACAACCTAACAGGATT
 K F N Y Y K C N L P L P I M S A I Y S R 11460
 TAAGTTTAATTATTACAAATGTAATTTACCATTGCCCATTTATGAGCGCAATATATTCTCG
 D F F R N E R F D I K L K I V A D A D W 11520
 TGATTTCTTCAGAAACGAACGTTTGTATTTAAATTAATAATTTGTTGCTGACGCTGATTG
 F L R C F I K W S K E K S P Y F I N D T 11580
 GTTCTCTGAGATGTTTCATCAAATGGAGTAAAGAGAAGTCACCTTATTTTATTAATGACAC
 T P I V R M G Y G G V S T D I S S Q V K 11640
 GACCCCTATTGTTAGAATGGGATATGGTGGGGTTTCGACTGATATTTCTTCTCAAGTTAA
 T T L E S F I V R K K N N I S C L N I Q 11700
 AACTACGCTAGAAAGTTTCATTGTACGCAAAAAGAATAATATATCCTGTTTAAACATACA
 L I L R Y A K I L V M V A I K N I F G N 11760
 GCTGATTCTTAGATATGCTAAATTTCTGGTGATGGTAGCGATCAAAAATATTTTGGCAA
 N V Y K L M H N G Y H S L K K I K N K I 11820
 TAATGTTTATAAATTAATGCATAACGGGTATCATTCCCTAAAGAAAATCAAGAATAAAAT

Start of orf11, End of orf10

M K I V Y I I T G L T C G G A E H L M T

*

ATGAAGATTGTTTATATAATAACCGGGCTTACTTGTGGTGGAGCCGAACACCTTATGACG 11880
 Q L A D Q M F I R G H D V N I I C L T G 11940
 CAGTTAGCAGACCAATGTTTATACGCGGGCATGATGTTAATATTATTTGTCTAACTGGT
 I S E V K P T Q N I N I H Y V N M D K N 12000
 ATATCTGAGGTAAAGCCAACACAAAATATTAATATTTCATTATGTTAATATGGATAAAAAT
 F R S F F R A L F Q V K K I I V A L K P 12060
 TTTAGAAGCTTTTTTAGAGCTTTATTTCAAGTAAAAAATAAATTGTCGCCTTAAAGCCA
 D I I H S H M F H A N I F S R F I R M L 12120
 GATATAATACATAGTCATATGTTTCATGCTAATATTTTATGTCGTTTTATTAGGATGCTG
 I P A V P L I C T A H N K N E G G N A R 12180
 ATCCAGCGGTGCCCCCTGATATGTACCCGACACAACAAAATGAAGGTGGCAATGCAAGG
 M F C Y R L S D F L A S I T T N V S K E 12240
 ATGTTTTGTATCGACTGAGTGATTTTTTAGCTTCTATTACTACAAATGTAAGTAAAGAG
 A V Q E F I A R K A T P K N K I V E I P 12300
 GCTGTTCAAGAGTTTATAGCAAGAAAGGCTACACCTAAAAATAAAATAGTAGAGATTCCG
 N F I N T N K F D F D I N V R K K T R D 12360
 AATTTTATTAATACAAATAAATTTGATTTTGATATTAATGTCAGAAAGAAAACGCGAGAT
 A F N L K D S T A V L L A V G R L V E A 12420
 GCTTTTAAATTTGAAAGACAGTACAGCAGTACTGCTCGCAGTAGGAAGACTTGTTGAAGCA
 K D Y P N L L N A I N H L I L S K T S N 12480
 AAAGACTATCCGAACCTTATTAATGCAATAAATCATTTGATTCTTTCAAAAACATCAAAT
 C N D F I L L I A G D G A L R N K L L D 12540
 TGTAATGATTTTATTTTGCTTATTGCTGGCGATGGCGCATTAAGAAATAAATTATTGGAT
 L V C Q L N L V D K V F F L G Q R S D I 12600
 TTGTTTGTCAATTGAATCTTGTGGATAAAGTTTCTTCTTGGGGCAAAGAAGTGATATT

15/96

K E L M C A A D L F V L S S E W E G F G 12660
 AAAGAATTAATGTGTGCTGCAGATCTTTTTGTTTTGAGTTCTGAGTGGGAAGGTTTTGGT
 L V V A E A M A C E R P V V A T D S G G 12720
 CTCGTTGTTGCAGAAGCTATGGCGTGTGAACGTCCCGTTGTTGCTACCGATTCTGGTGA
 V K E V V G P H N D V I P V S N H I L L 12780
 GTTAAAGAAGTCGTTGGACCTCATAATGATGTTATCCCTGTCAGTAATCATATTCTGTTG
 A E K I A E T L K I D D N A R K I I G M 12840
 GCAGAGAAAATCGCTGAGACACTTAAATAGATGATAACGCAAGAAAAATAATAGGTATG
 K N R E Y I V S N F S I K T I V S E W E 12900
 AAAAATAGAGAATATATTGTTTCCAATTTTTCAATTAAAACGATAGTGAGTGAGTGGGAG

End of orf11

R L Y F K Y S K R N N I I D * 12960
 CGCTTATATTTTAAATATTCCAAGCGTAATAATATAATTGAT TGAAAATATAAGTTTGTG
 CTCTGGATGCAATAGTTTCTCTATGCTGTTTTTTTTACTGGCTCCGTATTTTTACTTATAG 13020
 CTGGATTTTGTATATATCAGTATTAATCTGTCTCAACTTCATCTAGACTACATTCAAGC 13080

Start of gnd

M S K Q Q I
 CGCGCATGCGTCGCGCGGTGACTACACCTGACAGGAGTATGTAATGTCCAAGCAACAGAT 13140
 G V V G M A V M G R N L A L N I E S R G 13200
 CGGCGTCGTCGGTATGGCAGTGATGGGGCGCAACCTGGCGCTCAACATCGAAAGCCGCGG
 Y T V S I F N R S R E K T E E V V A E N 13260
 TTATACCGTCTCCATCTTCAACCGCTCCCGCGAGAAACTGAAGAAGTTGTTGCCGAGAA
 P D K K L V P Y Y T V K E F V E S L E T 13320
 CCCGGATAAGAACTGGTTCCTTATTACACGGTGAAAGAGTTTCGTCGAGTCTCTTGAAAC
 P R R I L L M V K A G A G T D A A I D S 13380
 CCCACGTCGTATCCTGTTAATGGTAAAAGCAGGGGCGGGAAGTATGCTGCTATCGATTCT
 L K P Y L D K G D I I I D G G N T F F Q 13440
 CCTGAAGCCGTATCTGGATAAAGGCGACATCATTATTGATGGTGGCAACACCTTCTTCCA
 D T I R R N R E L S A E G F N F I G T G 13500
 GGACACTATCCGTCGTAACCGTGAAGTGTCCGCGGAAGGCTTTAACTTCATCGGTACCGG
 V S G G E E G A L K G P S I M P G G Q K 13560
 CGTGTCCGGCGGTGAAGAGGGCGCCCTGAAAGGCCATCTATCATGCCAGGTGGCCAGAA
 E A Y E L V A P I L T K I A A V A E D G 13620
 AGAAGCGTATGAGCTGGTTGCGCCTATCCTGACCAAGATTGCTGCGGTTGCTGAAGATGG
 E P C I T Y I G A D G A G H Y V K M V H 13680
 CGAACCATGTATAACTTACATCGGTGCTGACGGTGCGGGTCACTACGTGAAGATGGTGCA
 N G I E Y G D M Q L I A E A Y S L L K G 13740
 CAACGGTATCGAATATGGCGATATGCAGCTGATTGCTGAAGCCTATTCTCTGCTTAAAGG
 G L N L S N E E L A T T F T E W N E G E 13800
 CGGCCTTAATCTGTCTAACGAAGAGCTGGCAACCACTTTTACCGAGTGGAATGAAGGCCA
 L S S Y L I D I T K D I F T K K D E E G 13860
 GCTAAGTAGTACCTGATTGACATCACCAAAGACATCTTCACCAAAAAAGATGAAGAGGG

16/96

K Y L V D V I L D E A A N K G T G K W T
TAAATACCTGGTTGATGTGATCCTGGACGAAGCTGCGAACAAAGGCACCGGTAAATGGAC 13920

S Q S S L D L G E P L S L I T E S V F A
CAGCCAGAGCTCTCTGGATCTGGGTGAACCGCTGTGCTGATCACCGAATCCGTATTTCGC 13980

R Y I S S L K D Q R I A A S K V L S G P
TCGCTACATCTCTTCTCTGAAAGACCAGCGCATTGCGGCATCTAAAGTGCTGTCTGGTCC 14040

Q A K L A G D K A E F V E K V R R A L Y
GCAGGCTAAACTGGCTGGTGATAAAGCAGAGTTTCGTTGAGAAAGTCCGTCGCGCGCTGTA 14100

L G K I V S Y A Q G F S Q L R A A S D E
CCTGGGTAAATCGTCTCTTATGCCCAAGGCTTCTCTCAACTGCGTGCCGCGTCTGACGA 14160

Y N W D L N Y G E I A K I F R A G C I I
ATACAACTGGGATCTGAACTACGGCGAAATCGCGAAGATCTTCCGCGCGGGCTGCATCAT 14220

R A Q F L Q K I T D A Y A E N K G I A N
TCGTGCGCAGTTTCCTGCAGAAAAATTACTGACGCGTATGCTGAAAACAAAGGCATTGCTAA 14280

L L L A P Y F K N I A D E Y Q Q A L R D
CCTGTTGCTGGCTCCGTACTTCAAAAATATCGCTGATGAATATCAGCAAGCGCTGCGTGA 14340

V V A Y A V Q N G I P V P T F S A A V A
TGTAAGTGGCTTATGCTGTGCAGAACGGTATTCCGGTACCGACCTTCTCTGCAGCGGTAGC 14400

Y Y D S Y R S A V L P A N L I Q A Q R D
CTACTACGACAGCTACCGTTCTGCGGTACTGCCGGCTAATCTGATTCAGGCACAGCGTGA 14460

Y F G A H T Y K R T D K E G V F H T G
TTACTTCGGTGCGCACACGTATAAACGCACTGATAAAGAAGGTGTGTTCCACACCG 14516

17/96

GTAACCAAGGGCGGTACGTGCATAAATTTTAATGCTTATCAAAACTATTAGCATTAAAAA 60

Start of orf1

M N K E T V S I I M P V Y N
TATATAAGAAATTCTCAAATGAAGAAAGAAACCGTTTCAATAATTATGCCCGTTTACAAT 120

G A K T I I S S V E S I I H Q S Y Q D F
GGGGCCAAAACATAATCTCATCAGTAGAATCAATTATACATCAATCTTATCAAGATTTT 180

V L Y I I D D C S T D D T F S L I N S R
GTTTTGTATATCATTGACGATTGTAGCACCAGATGATACATTTTCATTAATCAACAGTCGA 240

Y K N N Q K I R I L R N K T N L G V A E
TACAAAACAATCAGAAAATAAGAATATTGCGTAACAAGACAAAATTTAGGTGTTGCAGAA 300

S R N Y G I E M A T G K Y I S F C D A D
AGTCGAAATTATGGAATAGAAATGGCCACGGGGAAATATATTCTTTTGTGATGCGGAT 360

D L W H E K K L E R Q I E V L N N E C V
GATTTGTGGCAGAGAAAAAATTAGAGCGTCAAATCGAAGTGTTAAATAATGAATGTGTA 420

D V V C S N Y Y V I D N N R N I V G E V
GATGTGGTATGTTCTAATTATTATGTTATAGATAACAATAGAAATATTGTTGGCGAAGTT 480

N A P H V I N Y R K M L M K N Y I G N L
AATGCTCCTCATGTGATAAATTATAGAAAAATGCTCATGAAAACTACATAGGGAATTTG 540

T G I Y N A N K L G K F Y Q K K I G H E
ACAGGAATCTATAATGCCAACAAATTGGGTAAGTTTATCAAAAAAGATTGGTCACGAG 600

D Y L M W L E I I N K T N G A I C I Q D
GATTATTTGATGTGGCTGGAAATAATTAATAAAACAAATGGTGCTATTTGTATTCAAGAT 660

N L A Y Y M R S N N S L S G N K I K A A
AATCTGGCGTATTACATGCGTTCAAATAATTCATCTCGGGTAATAAAATTAAGCTGCA 720

K W T W S I Y R E H L H L S F P K T L Y
AAATGGACATGGAGTATATATAGAGAACATTTACATTTGTCCTTTCCAAAAACATTATAT 780

Y F L L Y A S N G V M K K I T H S L L R
TATTTTTTATTATATGCTTCAAATGGAGTCATGAAAAAATAACACATTCATCTATTAAGG 840

Start of orf2, End of orf1

R K E T K K *
V K S A A K L I F L F L F T
AGAAAGGAGACTAAAAAGTGAAGTCAGCGGCTAAGTTGATTTTTTTTATTCCTATTTACAC 900

L Y S L Q L Y G V I I D D R I T N F D T
TTTATAGTCTCCAGTTGTATGGGGTTATCATAGATGATCGTATAACAAATTTTGATACAA 960

K V L T S I I I I F Q I F F V L L F Y L
AGGTATTAAGTAGTATTATAATTATATTTTCAGATTTTTTTTGTATTATTATTTATCTAA 1020

T I I N E R K Q Q K K F I V N W E L K L
CGATTATAAATGAAAGAAAACAGCAGAAAAAATTTATCGTGAAGTGGGAGCTAAAGTTAA 1080

I L V F L F V T I E I A A V V L F L K E
TACTCGTTTTCTTTTTGTGACTATAGAAATTGCTGCTGTAGTTTTATTCTTAAAGAAG 1140

G I P I F D D D P G G A K L R I A E G N
GTATTCCTATATTTGATGATGATCCAGGGGGGGCTAAACTTAGAATAGCTGAAGGTAATG 1200

18/96

G L Y I R Y I K Y F G N I V V F A L I I 1260
 GACTTTACATTAGATATATTAAGTATTTTGGTAATATAGTTGTGTTGCATTAATTATTC
 L Y D E H K F K Q R T I I F V Y F T T I 1320
 TTTATGATGAGCATAAATTCAAACAGAGGACCATCATATTTGTATATTTTACAACGATTG
 A L F G Y R S E L V L L I L Q Y I L I T 1380
 CTTTATTTGGTTATCGTTCTGAATTGGTGTGCTCATTCTTCAATATATATTGATTACCA
 N I L S K D N R N P K I K R I I G Y F L 1440
 ATATCCTGTCAAAGGATAACCGTAATCCTAAAATAAAAAGAATAATAGGGTATTTTTTAT
 L V G V V C S L F Y L S L G Q D G E Q N 1500
 TGGTAGGGGTTGTATGCTCGTTGTTTATCTAAGTTTAGGACAAGACGGAGAACAAAATG
 D S Y N N M L R I I N R L T I E Q V E G 1560
 ACTCATATAATAATATGTTAAGGATAATTAATAGGTAAACAATAGAGCAAGTTGAAGGTG
 V P Y V V S E S I K N D F F P T P E L E 1620
 TTCCATATGTTGTTTCTGAATCTATTAAGAACGATTTCTTTCCGACACCAGAGTTAGAAA
 K E L K A I I N R I Q G I K H Q D L F Y 1680
 AGGAATTTAAAGCAATAATAATAGAATACAGGGAATAAAGCATCAAGACTTATTTTATG
 G E R L H K Q V F G D M G A N F L S V T 1740
 GAGAACGGTTACATAAACAAGTATTTGGAGACATGGGAGCAAATTTTTTATCAGTTACTA
 T Y G A E L L V F F G F L C V F I I P L 1800
 CGTATGGAGCAGAACTGTTAGTTTTTTTGGTTTTCTCTGTGTATTCTATCCCTTTAG
 G I Y I P F Y L L K R M K K T H S S I N 1860
 GGATATATATACCTTTTTTATCTTTTAAAGAGAATGAAAAAACCATAGCTCGATAAATT
 C A F Y S Y I I M I L L Q Y L V A G N A 1920
 GCGCATTCTATTCTATATATCATTATGATTTTATTGCAATACTTAGTGGCTGGGAATGCAT
 S A F F F G P F L S V L I M C T P L I L 1980
 CGGCCTTCTTTTTTGGTCCTTTTCTCTCCGTATTGATAATGTGTACTCCTCTGATCTTAT

Start of orf3

M K I S V I T V T Y
 L H D T L K R L S R N E N I S Y N C D L
 TGCATGATACGTTAAAGAGATTATCACGAAATGAAATATCAGTTATAACTGTGACTTAT 2040

End of orf2

N N A E G L E K T L S S L S I L K I K P
 *
 AATAATGCTGAAGGGTTAGAAAAAATTTAAGTAGTTTATCAATTTTAAAAATAAAACCT 2100
 F E I I I V D G G S T D G T N R V I S R 2160
 TTTGAGATTATTATAGTTGATGGCGGCTCTACAGATGGAACGAATCGTGTCTATTAGTAGA
 F T S M N I T H V Y E K D E G I Y D A M 2220
 TTTACTAGTATGAATATTACACATGTTTATGAAAAAGATGAAGGGATATATGATGCGATG
 N K G R M L A K G D L I H Y L N A G D S 2280
 AATAAGGGCCGAATGTTGGCCAAAGGCGACTTAATACATTATTTAAACGCCGCGATAGC
 V I G D I Y K N I K E P C L I K V G L F 2340
 GTAATTGGAGATATATATAAAATATCAAAGAGCCATGTTTGATTAAAGTTGGCCTTTTC
 E N D K L L G F S S I T H S N T G Y C H 2400
 GAAATGATAAACTTCTGGGATTTTCTTCTATAACCCATTCAAATACAGGGTATTGTTCAT

19/96

Q G V I F P K N H S E Y D L R Y K I C A 2460
 CAAGGGGTGATTTTCCCAAAGAATCATTGAGAATATGATCTAAGGTATAAAATATGTGCT
 D Y K L I Q E V F P E G L R S L S L I T 2520
 GATTATAAGCTTATTCAAGAGGTGTTTCTGGAAGGGTTAAGATCTCTATCTTTGATTACT
 S G Y V K Y D M G G V S S K K R I L R D 2580
 TCGGGTTATGTAAATATGATATGGGGGAGTATCTTCAAAAAAAGAATTTTAAGAGAT
 K E L A K I M F E K N K K N L I K F I P 2640
 AAAGAGCTTGCCAAAATTATGTTTGAAAAAATAAAAAAACCTTATTAAGTTTATTCCA
 I S I I K I L F P E R L R R V L R K M Q 2700
 ATTTCAATAATCAAAATTTTATTCCTGAACGTTTAAGAAGAGTATTGCGGAAAATGCAA

 Y I C L T L F F M K N S S P Y D N E * Start of orf4 End of orf3
 M I M N K I
 TATATTTGTCTAACTTTTATCTTCATGAAGAATAGTTCACCATATGATAATGAATAAAAT 2760
 K K I L K F C T L K K Y D T S S A L G R 2820
 CAAAAAATACTTAAATTTGCACTTTAAAAAATATGATACATCAAGTGCTTTAGGTAG
 E Q E R Y R I I S L S V I S S L I S K I 2880
 AGAACAGGAAAGGTACAGGATTATATCCTTGTCTGTTATTTCAAGTTTGATTAGTAAAT
 L S L L S L I L T V S L T L P Y L G Q E 2940
 ACTCTCACTACTTTCTCTTATATTAAGTAACTTTTACCTTATTTAGGACAAGA
 R F G V W M T I T S L G A A L T F L D L 3000
 GAGATTTGGTGTATGGATGACTATTACCAGTCTTGGTGCTGCTGACATTTTGGACTT
 G I G N A L T N R I A H S F A C G K N L 3060
 AGGTATAGGAAATGCATTAACAAACAGGATCGCACATTCAATTTGCGTGTGGCAAAAATTT
 K M S R Q I S G G L T L L A G L S F V I 3120
 AAAGATGAGTCGGCAAATTAGTGGTGGGCTCACTTTGCTGGCTGGATTATCGTTTGTGCT
 T A I C Y I T S G M I D W Q L V I K G I 3180
 AACTGCAATATGCTATATTACTTCTGGCATGATTGATTGGCAACTAGTAATAAAGGTAT
 N E N V Y A E L Q H S I K V F V I I F G 3240
 AAACGAGAATGTGTATGCAGAGTTACAACACTCAATTAAAGTCTTTGTAATCATATTTGG
 L G I Y S N G V Q K V Y M G I Q K A Y I 3300
 ACTTGGAATTTATTCAAATGGTGTGCAAAAAGTTTATATGGGAATACAAAAGCCTATAT
 S N I V N A I F I L L S I I T L V I S S 3360
 AAGTAATATTGTTAATGCCATATTTATATTGTTATCTATTATTACTCTAGTAATATCGTC
 K L H A G L P V L I V S T L G I Q Y I S 3420
 GAACTACATGCGGGACTACCAGTTTTAATTGTCAGCACTCTTGGTATTCAATACATATC
 G I Y L T I N L I I K R L I K F T K V N 3480
 GGGAACTCTATTTAACAATTAATCTTATTATAAAGCGATTAATAAAGTTTACAAAAGTTAA
 I H A K R E A P Y L I L N G F F F F I L 3540
 CATACATGCTAAAAAGAGAAGCTCCATATTTGATATTAAACGGTTTTTTCTTTTTTATTTT
 Q L G T L A T W S G D N F I I S I T L G 3600
 ACAGTTAGGCACTCTGGCAACATGGAGTGGTGATAACTTTATAATATCTATAACATTGGG

20/96

V T Y V A V F S I T Q R L F Q I S T V P 3660
 TGTTACTTATGTTGCTGTTTTTAGCATTACACAGAGATTATTTCAAATATCTACGGTCCC
 L T I Y N I P L W A A Y A D A H A R N D 3720
 TCTTACGATTTATAACATCCCGTTATGGGCTGCTTATGCAGATGCTCATGCACGCAATGA
 T Q F I K K T L R T S L K I V G I S S F 3780
 TACTCAATTTATAAAAAAGACGCTCAGAACATCATTGAAAATAGTGGGTATTTTCATCATT
 L L A F I L V V F G S E V V N I W T E G 3840
 CTTATTGGCCTTCATATTAGTAGTGTTCGGTAGTGAAGTCGTTAATATTTGGACAGAAGG
 K I Q V P R T F I I A Y A L W S V I D A 3900
 AAAGATTGAGGTACCTCGAACATTTCATAATAGCTTATGCTTTATGGTCTGTTATTGATGC
 F S N T F A S F L N G L N I V K Q Q M L 3960
 TTTTTCGAATACATTTGCAAGCTTTTTTAAATGGTTTGAACATAGTTAAACAACAAATGCT
 A V V T L I L I A I P A K Y I I V S H F 4020
 TGCTGTTGTAACATTGATATTGATCGCAATTCCAGCAAAATACATCATAGTTAGCCATTT
 G L T V M L Y C F I F I Y I V N Y F I W 4080
 TGGGTTAACTGTTATGTTGTACTGCTTCATTTTTATATATATTTGTAAATTACTTTATATG

Start of orf5, End of orf4

M K M

Y K C S F K K H I D R Q L N I R G *
 GTATAAATGTAGTTTTTAAAAACATATCGATAGACAGTTAAATATAAGAGGATGAAAATG 4140
 K Y I P V Y Q P S L T G K E K E Y V N E 4200
 AAATATATACCAGTTTACCAACCGTCATTGACAGGAAAAGAAAAGAATATGTAAATGAA
 C L D S T W I S S K G N Y I Q K F E N K 4260
 TGTCTGGAAGTCAACCGTGGATTTCATCAAAAGCAAACTATATTCAGAAAGTTTGAAAATAAA
 F A E Q N H V Q Y A T T V S N G T V A L 4320
 TTTGCGGAACAAAACCATGTGCAATATGCAACTACTGTAAAGTAATGGAACGGTTGCTCTT
 H L A L L A L G I S E G D E V I V P T L 4380
 CATTTAGCTTTTGTAGCGTTAGGTATATCGGAAGGAGATGAAGTTATTGTTCCAAACACTG
 T Y I A S V N A I K Y T G A T P I F V D 4440
 ACATATATAGCATCAGTTAATGCTATAAAATACACAGGAGCCACCCCCATTTCGTTGAT
 S D N E T W Q M S V S D I E Q K I T N K 4500
 TCAGATAATGAAACTTGGCAAAATGTCTGTTAGTGACATAGAACAAAATCACTAATAAA
 T K A I M C V H L Y G H P C D M E Q I V 4560
 ACTAAAGCTATTATGTGTGTCCTTTTATACGGACATCCATGTGATATGGAACAAATTGTA
 E L A K S R N L F V I E D C A E A F G S 4620
 GAAGTGGCCAAAAGTAGAAAATTTGTTTTGTAATTGAAGATTGCGCTGAAGCCTTTGCTTCT
 K Y K G K Y V G T F G D I S T F S F F G 4680
 AAATATAAAGCTAATATGTGCGAACATTTGGAGATATTTCTACTTTTACGTTTPTTGA
 N K T I T T G E G G M V V T N D K T L Y 4740
 AATAAACTATTACTACAGGTGAAGGTGGAATGTTGTCACGAATGACAAAACACTTTAT
 D R C L H F K G Q G L A V H R Q Y W H D 4800
 GACCGTTGTTTACATTTTAAAGCCCAAGCATTAGCTGTACATAGCCAAATATTGGCATGAC
 V I G Y N Y R M T N I C A A I G L A Q L 4860
 GTTATAGGCTACAAATTATAGGATGACAAATATCTGCGCTGCTATAGGATTAGCCCACTTA

21/96

E Q A D D F I S R K R E I A D I Y K K N GAACAAGCTGATGATTTTATATCAGGAAAACCTGAAATTGCTGATATTTATAAAAAAAT	4920
I N S L V Q V H K E S K D V F H T Y W M ATCAACAGTCTTTGTACAAAGTCCACAAGGAAAGTAAAGATGTTTTTCACACTTATTGGATG	4980
V S I L T R T A E E R E E L R N H L A D GTCTCAATTCTAACTAGGACCCGAGAGGAAAGAGAGGAATTAAGGAATCACCTTGCAGAT	5040
K L I E T R P V F Y P V H T M P M Y S E AAATCATCGAAACAAGGCGAGTTTTTTACCTGTCCACACGATGCGCAATGTACTCGGAA	5100
K Y Q K H P I A E D L G W R G I N L P S AAATATCAAAAGCACCTATAGCTGAGGATCTTGCTTGGCGTGGAAATTAATTTACCTAGT	5160
F P S L S N E Q V I Y I C E S I N E F Y TTCCCCAGCCTATCGAATGAGCAAGTTATTTATATTGTTGAATCTATTAACGAATTTTAT	5220
<div style="display: flex; justify-content: space-around; margin-top: 20px;"> <div style="text-align: center;"> <p>End of orf5</p> <p>S D K *</p> </div> <div style="text-align: center;"> <p>Start of orf6</p> <p>M K I A L N S D</p> </div> </div>	
AGTGATAAATAGCCTAAATATTTCTAAAGGTCATTCTGAAAAATTCGCTTGAATTCAGAT	5280
G F Y E W G G G I D F I K Y I L S I L E GGATTTTACGAGTGGGGCGGTGGAATTTGATTTTATTAAATATATTCTGTCAATATTAGAA	5340
T K P E I C I D I L L P R N D I H S L I ACGAAACCAGAAATATGTATCGATATTCTTTTACCGAGAAATGATATACATTCTCTTATA	5400
R E K A F P F K S I L K A I L K R E R P AGAGAAAAAGCATTTCCTTTTAAAAGTATATTTAAAGCAATTTTAAAGAGGGAAAGGCCT	5460
R W I S L N R F N E Q Y Y R D A F T Q N CGATGGATTTTCATTAAATAGATTTAATGAGCAATACTATAGAGATGCCTTTACACAAAAT	5520
N I E T N L T F I K S K S S A F Y S Y F AATATAGAGACGAATCTTACCTTTATTAAAAGTAAGAGCTCTGCCTTTTATTTCATATTTT	5580
D S S D C D V I L P C M R V P S G N L N GATAGTAGCGATTGTGATGTTATTCTTCCTTGCATGCGTGTTCCTTCGGGAAATTTGAAT	5640
K K A W I G Y I Y D F Q H C Y Y P S F F AAAAAAGCATGGATTGGTTATTTATGACTTTCAACACTGTTACTATCCTTCATTTTTT	5700
S K R E I D Q R N V F F K L M L N C A N AGTAAGCGAGAAATAGATCAAAGGAATGTGTTTTTTAAATTGATGCTCAATTGCGCTAAC	5760
N I I V N A H S V I T D A N K Y V G N Y AATATTATTGTTAATGCACATTCACTTATTACCGATGCAAATAAATATGTTGGGAATTAT	5820
S A K L H S L P F S P C P Q L K W F A D TCTGCAAAACTACATTCTCTTCCATTTAGTCCATGCCCTCAATTAAAAATGGTTCGCTGAT	5880
Y S G N I A K Y N I D K D Y F I I C N Q TACTCTGGTAATATTGCCAAATATAATATTGACAAGGATTATTTTATAATTTGCAATCAA	5940
F W K H K D H A T A F R A F K I Y T E Y TTTTGGAAACATAAAGATCATGCAACTGCTTTTAGGGCATTATAAAATTTTACTGAATAT	6000
N P D V Y L V C T G A T Q D Y R F P G Y AATCTGATGTTTATTAGTATGCACGGGAGCTACTCAAGATTATCGATTCCCTGGATAT	6060
F N E L M V L A K K L G I E S K I K I L TTTAATGAATTGATGGTTTTTGGCAAAAAGCTCGGAATTGAATCGAAAATTAAGATATTA	6120

22/96

G H I P K L E Q I E L I K N C I A V I Q
 GGGCATATACCTAAACTGAACAAATTGAATTAATCAAAAATTGCATTGCTGTAATACAA 6180
 P T L F E G G P G G G V T F D A I A L G
 CCAACCTTATTTGAAGCGGGCCTGGAGGGGGGTAACATTTGACGCTATTGCATTAGGG 6240
 K K V I L S D I D V N K E V N C G D V Y
 AAAAAAGTTATACTATCTGACATAGATGTCAATAAAGAAGTTAATTGCGGTGATGTATAT 6300
 F F Q A K N H Y S L N D A M V K A D E S
 TTCTTTCAGGCAAAAACCATTATTCATTAAATGACGCGATGGTAAAAGCTGATGAATCT 6360
 K I F Y E P T T L I E L G L K R R N A C
 AAAATTTTTTATGAACCTACAACTCTGATAGAATTGGGTCTCAAAAGACGCAATGCGTGT 6420
 A D F L L D V V K Q E I E S R S * **End of orf6**
 GCAGATTTTCTTTTAGATGTTGTGAAACAAGAAATTGAATCCCGATCT TAATATATTCAA 6480
Start of orf7
 M T K V A L I T G V T G Q D G S Y
 GAGGTATATAATGACTAAAGTCGCTCTTATTACAGGTGTAAGTGGACAAGATGGATCTTA 6540
 L A E F L L D K G Y E V H G I K R R A S
 TCTAGCTGAGTTTTTGTCTTGATAAAGGGTATGAAGTTCATGGTATCAAACGCCGAGCCTC 6600
 S F N T E R I D H I Y Q D P H G S N P N
 ATCTTTTAATACAGAACGCATAGACCATATTTATCAAGATCCACATGGTTCTAACCCTAAA 6660
 F H L H Y G D L T D S S N L T R I L K E
 TTTTCACTTGCATATGGAGATCTGACTGATTCATCTAACCTCACTAGAATTCTAAAGGA 6720
 V Q P D E V Y N L A A M S H V A V S F E
 GGTACAGCCAGATGAAGTATATAATTTAGCTGCTATGAGTCACGTAGCAGTTTCTTTTGA 6780
 S P E Y T A D V D A I G T L R L L E A I
 GTCTCCAGAATATACAGCCGATGTCGATGCAATTGGTACATTACGTTTACTGGAAGCAAT 6840
 R F L G L E N K T R F Y Q A S T S E L Y
 TCGCTTTTTTAGGATTGGAACAAACGCGTTTCTATCAAGCTTCAACCTCAGAATTATA 6900
 G L V Q E I P Q K E S T P F Y P R S P Y
 TGGACTTGTTTCAGGAAATCCCTCAAAAAGAATCCACCCCTTTTTATCCTCGTTCCCTTGA 6960
 A V A K L Y A Y W I T V N Y R E S Y G I
 TGCAGTTGCAAACTTTACGCATATTGGATCACGGTAAATTATCGAGAGTCATATGGTAT 7020
 Y A C N G I L F N H E S P R R G E T F V
 TTATGCATGTAATGGTATATTGTTCAATCATGAATCTCCACGCCGTGGAGAAACGTTTGT 7080
 T R K I T R G L A N I A Q G L E S C L Y
 AACAAGGAAAATTACTCGAGGACTTGCAAATATTGCACAAGGCTTGAATCATGTTTGTGA 7140
 L G N M D S L R D W G H A K D Y V R M Q
 TTAGGGAATATGGATTCTGTTACGAGATTGGGGACATGCAAAAGATTATGTTAGAATGCA 7200
 W L M L Q Q E Q P E D F V I A T G V Q Y
 ATGGTTGATGTTACAACAGGAGCAACCCGAAGATTTTGTGATTGCAACAGGAGTCCAATA 7260
 S V R Q F V E M A A A Q L G I K M S F V
 CTCAGTCCGTCAGTTTGTGCAATGGCAGCAGCACAACCTTGGTATTAAGATGAGCTTTGT 7320

23/96

G K G I E E K G I V D S V E G Q D A P G 7380
 TGGTAAAGGAATCGAAGAAAAAGGCATTGTAGATTCCGTTGAAGGACAGGATGCTCCAGG
 V K P G D V I V A V D P R Y F R P A E V 7440
 TGTGAAACCAGGTGATGTCATTGTTGCTGTTGATCCTCGTTATTTCGACCAGCTGAAGT
 D T L L G D P S K A N L K L G W R P E I 7500
 TGATACTTTGCTTGGAGATCCGAGCAAAGCTAATCTCAAAC TTGGTTGGAGACCAGAAAT
 T L A E M I S E M V A K D L E A A K K H 7560
 TACTCTTGCTGAAATGATTTCTGAAATGGTTGCCAAAGATCTTGAAGCCGCTAAAAACA

Start of orf8, End of orf7

M M M N K
 S L L K S H G F S V S L A L E *
 TTCTCTTTTAAATCGCATGGTTTTTCTGTAAGCTTAGCTCTGGAATGATGATGAATAAG 7620
 Q R I F I A G H Q G M V G S A I T R R L 7680
 CAACGTATTTTATTGCTGGTCACCAAGGAATGGTTGGATCAGCTATTACCCGACGCCTC
 K Q R D D V E L V L R T R D E L N L L D 7740
 AAACAACGTGATGATGTTGAGTTGGTTTTACGTACTCGGGATGAATTGAACTTGTGGAT
 S S A V L D F F S S Q K I D Q V Y L A A 7800
 AGTAGCGCTGTTTTGGATTTTTTTCTTCACAGAAAATCGACCAGGTTTATTTGGCAGCA
 A K V G G I L A N S S Y P A D F I Y E N 7860
 GCAAAAGTCGGAGGTATTTTAGCTAACAGTTCTTATCCTGCCGATTTTATATATGAGAAT
 I M I E A N V I H A A H K N N V N K L L 7920
 ATAATGATAGAGGCGAATGTCATTCATGCTGCCACAAAAATAATGTAAATAAACTGCTT
 F L G S S C I Y P K L A H Q P I M E D E 7980
 TTCCTCGGTTCTGCTGTATTTATCCTAAGTTAGCACACCAACCGATTATGGAAGACGAA
 L L Q G K L E P T N E P Y A I A K I A G 8040
 TTATTACAAGGGAACTTGAGCCAACAAATGAACCTTATGCTATCGCAAAAATTGCAGGT
 I K L C E S Y N R Q F G R D Y R S V M P 8100
 ATTAATTATGTGAATCTTATAACCGTCAGTTTGGGCGTGATTACCGTTCAGTAATGCCA
 T N L Y G P N D N F H P S N S H V I P A 8160
 ACCAATCTTTATGGTCCAAATGACAATTTTCATCCAAGTAATTCTCATGTGATTCCGGCG
 L L R R F H D A V E N N S P N V V V W G 8220
 CTTTTGCGCCGCTTTCATGATGCTGTGGAAAACAATTCCTCGAATGTTGTTGTTGGGGA
 S G T P K R E F L H V D D M A S A S I Y 8280
 AGTGGTACTCCAAAGCGTGAATTCTTACATGTAGATGATATGGCTTCTGCAAGCATTAT
 V M E M P Y D I W Q K N T K V M L S H I 8340
 GTCATGGAGATGCCATACGATATATGGCAAAAAATACTAAAGTAATGTTGTCTCATATC
 N I G T G I D C T I C E L A E T I A K V 8400
 AATATTGGAACAGGTATTGACTGCACGATTTGTGAGCTTGCGGAAACAATAGCAAAAGTT
 V G Y K G H I T F D T T K P D G A P R K 8460
 GTAGGTTATAAAGGGCATATTACGTTTCGATACAACAAAGCCGATGGAGCCCCCTCGAAAA
 L L D V T L L H Q L G W N H K I T L H K 8520
 CTACTTGATGTAACGCTTCTTCATCAACTAGGTTGGAATCATAAAATTACCCTTCACAAG

24/96

End of orf8

G L E N T Y N W F L E N Q L Q Y R G *
GGTCTTGAATAACATACAACTGGTTTCTTGAAAACCAACTTCAATATCGGGGG TAATAA 8580

Start of orf9

M F L H S Q D F A T I V R S T P L I S I
TGTTTTACATTTCCCAAGACTTTGCCACAATTGTAAGGTCTACTCCTCTTATTTCTATAG 8640

D L I V E N E F G E I L L G K R I N R P
ATTTGATTGTGGAACGAGTTTGGCGAAATTTGCTAGGAAAACGAATCAACCGCCCGG 8700

A Q G Y W F V P G G R V L K D E K L Q T
CACAGGGCTATTGGTTCGTTCTGCTGGTGGTAGGTGTTGAAAGATGAAAAATTGCAGACAG 8760

A F E R L T E I E L G I R L P L S V G K
CCTTTGAACGATTGACAGAAATTGAACTAGGAATTCGTTTGCCTCTCTCTGTGGGTAAGT 8820

F Y G I W Q H F Y E D N S M G G D F S T
TTTATGGTATCTGGCAGCACTTCTACGAAGACAATAGTATGGGGGGAGACTTTTCAACGC 8880

H Y I V I A F L L K L Q P N I L K L P K
ATTATATAGTTATAGCATTCTTCTTAAATTACAACCAAACATTTTGAAATTACCGAAGT 8940

S Q H N A Y C W L S R A K L I N D D D V
CACAACATAATGCTTATTGCTGGCTATCGCGAGCAAAGCTGATAAATGATGACGATGTGC 9000

H Y N C R A Y F N N K T N D A I G L D N
ATTATAATTGTCGCGCATATTTTAAACAATAAAACAAATGATGCGATTGGCTTAGATAATA 9060

Start of orf10 End of orf9

M S D A P I I A V V M A G G T G S
K D I I C L M R Q *
AGGATATAATATGTCTGATGCGCCAATAATTGCTGTAGTTATGGCCGGTGGTACAGGCAG 9120

R L W P L S R E L Y P K Q F L Q L S G D
TCGTCTTTGGCCACTTTCTCGTGAACATATCCAAAGCAGTTTTTACAACCTCTCTGGTGA 9180

N T L L Q T T L L R L S G L S C Q K P L
TAACACCTTGTTACAAACGACTTTGCTACGACTTTTCAGGCCTATCATGTCAAAAACCAT 9240

V I T N E Q H R F V V A E Q L R E I N K
AGTGATAACAAATGAACAGCATCGCTTTGTTGTGGCTGAACAGTTAAGGGAAATAAATAA 9300

L N G N I I L E P C G R N T A P A I A I
ATTAATGGTAATATTATTCTAGAACCATGCGGGCGAAATACTGCACCAGCAATAGCGAT 9360

S A F H A L K R N P Q E D P L L L V L A
ATCTGCGTTTCATGCGTTAAACGTAATCCTCAGGAAGATCCATTGCTTCTAGTTCTTGC 9420

A D H V I A K E S V F C D A I K N A T P
GGCAGACCACGTTATAGCTAAAGAAAGTGTTTCTGTGATGCTATTAAAAATGCAACTCC 9480

I A N Q G K I V T F G I I P E Y A E T G
CATCGTAATCAAGGTAATAATTGTAACGTTTGAATTATACCAGAATATGCTGAAACTGG 9540

Y G Y I E R G E L S V P L Q G H E N T G
TTATGGGTATATTGAGAGAGGTGAACATCTGTACCGCTTCAAGGGCATGAAAATACTGG 9600

F Y Y V N K F V E K P N R E T A E L Y M
TTTTTATTATGTAAATAAGTTTGTGCGAAAAGCCTAATCGTGAAACCGCAGAATTGTATAT 9660

T S G N H Y W N S G I F M F K A S V Y L
GACTTCTGGTAATCACTATTGGAATAGTGGAATATTTCATGTTTAAGGCATCTGTTTATCT 9720

25/96

E E L R K F R P D I Y N V C E Q V A S S 9780
 TGAGGAATTGAGAAAATTTAGACCTGACATTTACAATGTTTGTGAACAGGTTGCCTCATC
 S Y I D L D F I R L S K E Q F Q D C P A 9840
 CTCATACATTGATCTAGATTTTATTTCGATTATCAAAAGAACAATTTCAAGATTGTCCTGC
 E S I D F A V M E K T E K C V V C P V D 9900
 TGAATCTATTGATTTTGCTGTAATGGAAAAACAGAAAAATGTGTTGTATGCCCTGTTGA
 I G W S D V G S W Q S L W D I S L K S K 9960
 TATTGGTTGGAGTGACGTTGGATCTTGGCAATCGTTATGGGACATTAGTCTAAAATCGAA
 T G D V C K G D I L T Y D T K N N Y I Y 10020
 AACAGGAGATGTATGTAAAGGTGATATATTAACCTATGATACTAAGAATAATTATATCTA
 S E S A L V A A I G I E D M V I V Q T K 10080
 CTCTGAGTCAGCGTTGGTAGCCGCCATTGGAATTGAAGATATGGTTATCGTGCAAACCTAA
 D A V L V S K K S D V Q H V K K I V E M 10140
 AGATGCCGTTCTTGTGTCTAAAAAGAGTGATGTACAGCATGTAAAAAAAATAGTCGAAAT
 L K L Q Q R T E Y I S H R E V F R P W G 10200
 GCTTAAATTGCAGCAACGTACAGAGTATATTAGTCATCGTGAAGTTTCCGACCATGGGG
 K F D S I D Q G E R Y K V K K I I V K P 10260
 AAAATTTGATTGCGATTGACCAAGGTGAGCGATACAAAGTCAAGAAAAATTATTGTGAAACC
 G E G L S L R M H H H R S E H W I V L S 10320
 TGGTGAGGGGCTTTCTTTAAGGATGCATCACCATCGTTCTGAACATTGGATCGTGCTTTC
 G T A K V T L G D K T K L V T A N E S I 10380
 TGGTACAGCAAAAGTAACCTTGGCGATAAACTAACTAGTCACCGCAAATGAATCGAT
 Y I P L G A A Y S L E N P G I I P L N L 10440
 ATACATTCCCTTGGCGCAGCGTATAGTCTTGAGAATCCGGGCATAATCCCTCTTAATCT
 I E V S S G D Y L G E D D I I R Q K E R 10500
 TATTGAAGTCAGTTCAAGGGATTATTGGGAGAGGATGATATTATAAGACAGAAAGAACG
 End of orf10 Start of orf11
 Y K H E D * M K S L T C F K A Y D I R 10560
 TTACAACATGAAGATTAACATATGAAATCTTTAACCTGCTTTAAAGCCTATGATATTGC
 G K L G E E L N E D I A W R I G R A Y G 10620
 CGGGAAATTAGGCGAAGAACTGAATGAAGATATTGCCTGGCGCATGGGCGTGCCTATGG
 E F L K P K T I V L G G D V R L T S E A 10680
 CGAATTTCTCAAACCGAAAACCATTTGTTTTAGGCGGTGATGTCCGCCTCACCAGCGAAGC
 L K L A L A K G L Q D A G V D V L D I G 10740
 GTTAAACTGGCGCTTGCGAAAGGTTTACAGGATGCGGGCGTCGATGTGCTGGATATCGG
 M S G T E E I Y F A T F H L G V D G G I 10800
 TATGTCCGGCACCGAAGAGATCTATTTGCCACGTTCCATCTCGGACTGGATGGCGGCAT
 E V T A S H N P M D Y N G M K L V R E G 10860
 CGAAGTTACCGCCAGCCATAACCCGATGGATTACAACGGCATGAAGCTGGTGCGCGAAGG
 A R P I S G D T G L R D V Q R L A E A N 10920
 GGCTCGCCCGATCAGCGGTGATACCGGACTGCGCGATGTCCAGCGTCTGCCAGAAGCCAA
 D F P P V D E T K R G R Y Q Q I N L R D 10980
 TGACTTCCCTCCTGTCGATGAAACCAAACGTGGTTCGCTATCAGCAAATCAATCTGCGTGA

26/96

A Y V D H L F G Y I N V K N L T P L K L
 CGCTTACGTTGATCACCTGTTCCGGTTATATCAACGTCAAAAACCTCACGCCGCTCAAGCT 11040
 V I N S G N G A A G P V V D A I E A R F
 GGTGATCAACTCCGGGAACGGCGCAGCGGGTCCGGTGGTGGACGCCATTGAAGCCCCGATT 11100
 K A L G A P V E L I K V H N T P D G N F
 TAAAGCCCTCGGCGCACCGGTGGAATTAATCAAAGTACACAACACGCCGACGGCAATTT 11160
 P N G I P N P L L P E C R D D T R N A V
 CCCCACGGTATTCTTAACCCGCTGCTGCCGAATGCCGCGACGACACCCGTAATGCGGT 11220
 I K H G A D M G I A F D G D F D R C F L
 CATCAAACACGGCGCGGATATGGGCATTGCCCTTTGATGGCGATTTTGACCGCTGTTTCCT 11280
 F D E K G Q F I E G Y Y I V G L L A E A
 GTTTGACGAAAAAGGGCAGTTTATCGAGGGCTACTACATTGTTCGGCCTGCTGGCAGAAGC 11340
 F L E K N P G A K I I H D P R L S W N T
 GTTCCTCGAAAAAATCCCGGCGCGAAGATCATCCACGATCCACGTCTCTCCTGGAACAC 11400
 V D V V T A A G G T P V M S K T G H A F
 CGTTGATGTGGTGACTGCCCGAGGCGGCACCCCGTAATGTCGAAAACCGGACACGCCTT 11460
 I K E R M R K E D A I Y G G E M S A H H
 TATTAAAGAACGTATGCGCAAGGAAGACGCCATCTACGGTGGCGAAATGAGCGCTCACCA 11520
 Y F R D F A Y C D S G M I P W L L V A E
 TTACTTCCGTGATTTTCGCTTACTGCGACAGCGGCATGATCCCGTGGCTGCTGGTCGCCGA 11580
 L V C L K G K T L G E M V R D R M A A F
 ACTGGTGTGCCTGAAAGGAAAAACGCTGGGCGAAATGGTGC GCGACCGGATGGCGGCGTT 11640
 P A S G E I N S K L A Q P V E A I N R V
 TCCGGAAGCGGTGAGATCAACAGCAAACCTGGCGCAACCCGTTGAGGCAATTAATCGCGT 11700
 E Q H F S R E A L A V D R T D G I S M T
 GGAACAGCATTTTAGCCGCGAGGCGCTGGCGGTGGATCGCACCGATGGCATCAGCATGAC 11760
 F A D W R F N L R S S N T E P V V R L N
 CTTTGCCGACTGGCGCTTTAACCTGCGCTCCTCAACACCGAACCAGGTGGTGGCGTTGAA 11820
 V E S R G D V K L M E K K T K A L L K L
 TGTGGAATCACGCGGTGATGTAAAGCTAATGGAAAAGAAAATAAAGCTCTTCTTAAATT 11880
End of orf11
 L S E *
 GCTAAGTGAGTGATTATTTACATTAATCATTAAGCGTATTTAAGATTATATTAAAGTAAT 11940
 GTTATTGCGGTATATGATGAATATGTGGGCTTTTATGTATAACGACTATACCGCAACT 12000
Start of H-repeat
 TTATCTAGGAAAAGATTAATAGAAATAAAGTTTTGTACTGACCAATTTGCATTTTCACGTC 12060
 ACGATTGAGACGTTCCTTTGCTTAAGACATTTTTTCATCGCTTATGTAATAACAAATGTG 12120
 CCTTATATAAAAAGGAGAACAAAATGGAACCTAAAATAATTGAGACAATAGATTTTTTATT 12180
 ATCCCTGTTTACGATATTATAGCCAAAGTTGTATCCTGCATCAGTCCTGCAATATTTTAC 12240
 GAGTGCTTTGTAACTGAATACATGTCTGCCATTTTCCAGATGATAACGACGTCATCGCA 12300
 ATTGATGGTAAAACACTTCGGCACACTTATGACAAGAGTCGTCGAGAGGAGTGTTTCAT 12360

27/96

GTCATTAGTGCCTTTCAGCAATGCACAGTCTGGTCCCTCGGATAGATCAAGACGGATGAGA 12420
AACCTAATGCGTTCACAGTTATTCATGAACTTCTAAAATGATGGGTATTAAAGGAAAAA 12480
TAATCATAACTGATGCGATGGCTTGCCAGAAAGATATTGCAGAGAAGATATAAAAAACAGA 12540
GATGTGATTATTTATTCGCTGTAAAAGGAAATAAGAGTCGGCTTAATAGAGTCTTTGAGG 12600
AGATATTTACGCTGAAAGAATTAAATAATCCAAAACATGACAGTTACGCAATTAGTGAAA 12660
AGAGGCACGGCAGAGACGATGTCCGTCTTCATATTGTTTGAGATGCTCCTGATGAGCTTA 12720
TTGATTTACGCTTTGAATGGAAAGGGCTGCAGAAATTTATGAATGGCAGTCCACTTTCTCT 12780
CAATAATAGCAGAGCAAAAGAAAGAATCCGAAATGACGATCAAATATTATATTAGATCTG 12840
CTGCTTTAACCGCAGAGAAGTTGCCACAGTAAATCGAAATCACTGGCGCATGGAGAATA 12900
AGTTGCACAGTAGCCTGATGTGGTAATGAATGAAATCGACTATAATATAAGAAGGCGAGT 12960
TGCATTGCAATGATTTTCTAGAATGCGGCACATCGCTATTAATATCTGACAATGATAATG 13020
TATTCAAGGCAGGATTATCATGTAAGATGCGAAAAGCAGTCATGGACAGAAACTTCCTAG 13080

End of the H-repeat

CGTCAGGCATTGCAGCGTGCAGGGCTTTCATAATCTTGCAT TGGTTTGTGATAAGATATTTT 13140

Start of orf12

M N L Y G I F G A G S Y G R E
TTTGGAGATGGGAAAATGAATTTGTATGGTATTTTTTGGTGCCTGGAAGTTATGGTAGAGAA 13200
T I P I L N Q Q I K Q E C G S D Y A L V
ACAATACCCATTCTAAATCAACAAATAAAGCAAGAATGTGGTTCTGACTATGCTCTGGTT 13260
F V D D V L A G K K V N G F E V L S T N
TTTGTGGATGATGTTTTGGCAGGAAAGAAAGTTAATGGTTTTGAAGTGCTTCAACCAAC 13320
C F L K A P Y L K K Y F N V A I A N D K
TGCTTTCTAAAAGCCCTTATTTAAAAAAGTATTTTAATGTTGCTATTGCTAATGATAAG 13380
I R Q R V S E S I L L H G V E P I T I K
ATACGACAGAGAGTGTCTGAGTCAATATTATTACACGGGGTTGAACCAATAACTATAAAA 13440
H P N S V V Y D H T M I G S G A I I S P
CATCCAAATAGCGTTGTTTATGATCATACTATGATAGGTAGTGGCGCTATTATTTCTCCC 13500
F V T I S T N T H I G R F F H A N I Y S
TTTGTACAATATCTACTAATACTCATATAGGGAGGTTTTTTCATGCAAACATATACTCA 13560
Y V A H D C Q I G D Y V T F A P G A K C
TACGTTGCACATGATTGTCAAATAGGAGACTATGTTACATTTGCTCCTGGGGCTAAATGT 13620
N G Y V V I E D N A Y I G S G A V I K Q
AATGGATATGTTGTTATTGAAGACAATGCATATATAGGCTCGGGTGCAGTAATTAAGCAG 13680
G V P N R P L I I G A G A I I G M G A V
GGTGTTCCTAATCGCCCACTTATTATTGGCGCGGGAGCCATTATAGGTATGGGGGCTGTT 13740
V T K S V P A G I T V C G N P A R E M K
GTCATAAAAGTGTTCCTGCCGGTATAACTGTGTGCGGAAATCCAGCAAGAGAAATGAAA 13800

End of orf12

R S P T S I *
AGATCGCCAACATCTATT TAATGGGAATGCGAAAACACGTTCCAAATGGGACTAATGTTT 13860

WO 99/61458

PCT/AU99/00385

28/96

AAAATATATATAATTTTCGCTAATTTACTAAATTATGGCTTCTTTTAAAGCTATCCTTTAC 13920

TTAGTTATTACTGATACAGCATGAAATTTATAATACTCTGATACATTTTATACGTTATT 13980

CAAGCCGCATATCTAGCGGTAACCCCTGACAGGAGTAAACAATG 14024

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2

29/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATAATATCAACAAG
AACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGCGTATTAACAGC
GCGAAGGATGACGCCGCGGGTCAGGCGATTGCTAACCGTTTTACTTCTAACATTAAAGGC
CTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCTGTTGCACAGACCACTGAAGGC
GCGCTGTCCGAAATCAACAACAACCTTACAGCGTATCCGTGAGCTGACGGTTCAGGCTTCT
ACCGGGACTAACTCTGATTCGGATCTGGACTCCATTTCAGGACGAAATCAAATCCCGTCTC
GACGAAATTGACCGCGTATCCGGTCAGACCCAGTTCAACGGCGTGAACGTACTGGCAAAA
GACGGTTCGATGAAAATTCAGGTAGGTGCGAACGACGGCCAGACTATCACTATTGATCTG
AAGAAAATTGACTCTGATACGCTGGGGCTGAATGGTTTTAACGTGAATGGTTCGGGTACG
ATAGCCAATAAAGCGGCGACCATTAGCGACCTGACAGCAGCGAAAATGGATGCTGCAACT
AATACTATAACTACAACAAATAATGCGCTGACTGCATCAAAGGCCCTTGATCAACTGAAA
GATGGTGACACTGTTACTATCAAAGCAGATGCAGCTCAAACGCCAGGTCTATACATAC
AATGCATCTGCTGGTAACCTTCTCATTAGTAATGTATCGAATAATACTTCAGCAAAAGCA
GGTGATGTAGCAGCTAGCCTTCTCCCGCCGGCTGGGCAAACTGCTAGTGGTGTTTACAAA
GCAGCAAGCGGTGAAGTGAACCTTGTATGTTGATGCGAATGGTAAAATTACAATCGGAGGA
CAGGAAGCCTATTTAACTAGTGATGGTAACTTAACTACAAACGATGCTGGTGGTGCGACT
GCGGCTACGCTTGATGGTTTTATTCAAGAAAGCTGGTGATGGTCAATCAATCGGGTTAAT
AAGACTGCATCAGTCACGATGGGGGGAACAACTTATAACTTTAAAACGGGTGCTGATGCT
GGTGCTGCAACTGCTAACGCAGGGGTATCGTTCACTGATACAGCTAGCAAAGAAACCGTT
TTAAATAAAGTGGCTACAGCTAAACAAGGCACAGCAGTTGCAGCTAACGGTGATACATCC
GCAACAATTACCTATAAATCTGGCGTTCAGACGTATCAGGCGGTATTTGCCGCAGGTGAC
GGTACTGCTAGCGCAAAATATGCCGATAATACTGACGTTTCTAATGCAACAGCAACATAC
ACAGATGCTGATGGTGAAATGACTACAATTGGTTCATACACCACGAAGTATTCAATCGAT
GCTAACAACGGCAAGGTAACCTGTTGATTCTGGAACCTGGTTCGGGTAAATATGCGCCGAAA
GTCGGGGCTGAAGTATATGTTAGTGCTAATGGTACTTTAACAACAGATGCAACTAGCGAA
GGCACAGTAACAAAAGATCCACTGAAAGCTCTGGATGAAGCTATCAGCTCCATCGACAAA
TTCCGTTTCATCCCTGGGGGCTATCCAAAACCGTTTGGATTCCGCCGTACCAACCTGAAC
AACACCACTACCAACCTGTCTGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACC
GAAGTGTCACATGTGAAAGCGCAGATTATCCAGCAGGCCGGTAACTCCGTGCTGGCA
AAAGCCAACCAGGTACCGCAGCAGGTTCTGTCTCTACTGCAGGGTTAA

Figure 7

30/96

AACAAATCTCAGTCTTCTCTTAGCTCTGCTATT
GAGCGTCTGTCTTCTGGTCTGCGTATTAAACAGCGCAAAGACGATGCAGCAGGTCAGGCG
ATTGCTAACCGTTTTACGGCAAATATTAAAGGTCTGACCCAGGCTTCCCGTAACGCAAAT
GATGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTG
CAGCGTATTCGTGAACTTTCTGTTCAAGGCAACTAACGGTACTAACTCTGACAGTGACCTG
ACCTCCATCCAGTCCGAAATCCAGCAGCGTCTGAGTGAAATTGACCGTGTTTCTGGTCAG
ACTCAGTTTAAACGGCGTTAAAGTGCTGGCTTCTGATCAGGATATGACTATTCAGGTTGGT
GCAAACGACGGCGAAACAATTACTATTAACTGCAGGAAATTAATCCGACACACTGGGA
TTATCTGGTTTTGGTATTAAAGATCCTACTAAATTTAAAGCCGCAACGGCTGAAACAACC
TATTTTGGATCGACAGTTAAGCTTGCTGACGCTAATACACTTGATGCAGATATTACAGCT
ACAGTTAAAGGCACTACGACTCCGGGCCAACGTGACGGTAATATTATGTCTGATGCTAAC
GGTAAGTTGTACGTTAAAGTTGCCGGTTCAGATAAACCCGCTGAAAATGGTTATTATGAA
GTTACTGTGGAGGATGATCCGACATCTCCTGATGCAGGTAAGCTGAAGCTGGGGGCTCTA
GCGGGTACCCAGCCTCAAGCTGGTAATTTAAAGGAAGTCACAACGGTGAAAGGGAAGGGG
GCTATTGATGTTTCAGTTGGGTACTGATACCGCAACCGCTTCTATCACAGGTGCAAACTC
TTTAAGTTAGAAGACGCCAATGGCAAAGATACTGGTTCATTTGCGTTGATTGGTGATGAC
GGTAAACAGTATGCAGCGAATGTTGATCAGAAAACAGGAGCAGTTTCCGTTAAAACAATG
TCTTACACTGATGCTGACGGTGTCAAACACGACAATGTTAAAGTTGAACTGGGTGGAAGC
GATGGCAAAACCGAAGTTGTAAGTGAACCGATGGCAAACTTACAGTGTTAGTGATTTA
CAAGGTAAGAGCCTGAAAAGTATTCTATTGCAGCAATTTCTACGCAGAAAACAGAAGAT
CCTTTGGCTGCTATCGATAAAGCACTGTCTCAGGTTGACTCGTTGCGTTCTAACCTAGGT
GCAATTCAAAATCGTTTTGACTCTGCCATACCAACCTTGGCAACACCGTAAACAACCTG
TCTTCTGCCCGTAGCCGTATCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCT
CGTGCGCAGATCCTGCAACAAGCGGGTACCTCTGTTCTGGCGCAG

Figure 8

31/96

AACAAATCTCAGTCTTCTCTGAGCTCCGCCATTGAACGTCTCTCTTC
TGGCCTGCGTATTAAACAGTGCTAAAGATGACGCAGCAGGTGAGGCGATTGCTAACCGTTT
TACAGCAAATATTAAAGGTCTGACTCAGGCTTCCCGTAACGCGAATGATGGTATTTCTGT
TGCGCAGACCACTGAAGGTGCGCTTTCTGAAATCAACAATAACTTACAGCGTATTCGTGA
ATTGTCAGTACAGGCCACTAATGGTACAACTCTGACTCCGACCTGAATTCAATTCAGGA
TGAAATTACACAACGCCTTAGTGAAATTGATCGTGTTTCTAACCAGACACAATTTAATGG
TGTAAGTTCTGGCTTCTGATCAGACTATGAAAATTCAAGTAGGTGCGAACGATGGTGA
AACCATTGAGATTGCCCTTGATAAAATTGATGCTAAAACCTTGGGGCTTGATAACTTTAG
CGTAGCACCAGGAAAAGTTCCAATGTCCTCTGCGGTTGCACTTAAGAGCGAAGCCGCTCC
TGACTTAACCTAAGGTAAATGCAACTGATGGTAGTGTGGGAGGTGCTAAAGCATTCCGTAG
CAATTATAAAATGCTGATGTTGAACTTATTTTGGTACCGGTAATGTACAAGATACAAA
GGATACAACCTGATGCGACCGGTACTGCAGGAACAAAAGTTTATCAAGTACAGGTGGAAGG
GCAGACTTATTTTGGTGGTCAAGATAATAATACCAACACGAACGGTTTTACATTATTGAA
ACAAAACCTCTACAGGTTATGAAAAAGTTTCAAGGTGGGTGGTAAGGATGTTTCAAGTTAGCAA
CTTTGGTGGTTCGTGTAAGTGCATTTGTTGAAGATAATGGTTCTGCCACATCAGTTGATTT
AGCTGCGGGTAAAATGGGTAAAGCATTAGCTTATAATGATGCACCAATGTCTGTTTATTT
TGGGGGAAAAAACCTAGATGTCCACCAAGTACAAGATACCCAAGGGAATCCTGTACCTAA
TTCATTTGCTGCTAAAACATCAGACGGCACCTACATTGCAGTAAATGTAGATGCCGCTAC
AGGTAACACGTCTGTTATTACTGATCCTAATGGTAAGGCAGTTGAATGGGCAGTAAAAAA
TGATGGTTCTGCACAGGCAATTATGCGTGAAGATGATAAGGTTTATACAGCCAATATCAC
GAATAAGACGGCAACCAAAGGTGCTGAACTCAGTGCCTCAGATTTGAAAGCCTTAGCAAC
CACAAATCCATTATCCACATTAGACGAAGCTTTGGCAAAAAGTTGATAAGTTGCGCAGTTC
TTTGGGTGCAGTACAAAACCGTTTCGACTCTGCCATCACCAACCTTGGCAACACCGTAAA
CAACCTGTCTTCTGCCCCGTAGCCGTATAGAAGATGCTGACTACGCAACCGAAGTGTCTAA
CATGTCTCGTGCGCAGATCCTGCAACAAGCGGGTACCTCTGTTCTGGCACAG

Figure 9

32/96

AACAAAAACCAGTCTGCGCTGTCGACTTCTATCGAG
CGCCTCTCTTCTGGTCTGCGTATTAACAGCGCTAAAGATGACGCCGCGGGCCAGGCGATT
GCTAACCGCTTTACTTCTAACATCAAAGGTCTGACTCAGGCCGCACGTAACGCCAACGAC
GGTATTTCTCTGGCGCAGACGGCTGAAGGCGCGCTGTCAGAGATTAACAACAACCTTGCAG
CGTATTCGTGAACTGACCGTTCAGGCCTCTACCGGCACGAACTCTGATTCCGACCTGTCT
TCTATTCAGGACGAAATCAAATCCCGTCTTGATGAAATTGACCGTGTATCTGGTCAGACC
CAGTTCAACGGTGTGAACGTGCTGTCGAAAAACGATTTCGATGAAGATTGAGATTGGTGCC
AATGATAACCAGACGATCAGCATTGGCTTGCAACAAATCGACAGTACCACTTTGAATCTG
AAAGGATTTACCGTGTCCGGCATGGCGGATTTTCAGCGCGGCGAACTGACGGCTGCTGAT
GGTACAGCAATTGCTGCTGCGGATGTCAAGGATGCTGGGGGTAAACAAGTCAATTTACTG
TCTTACACTGACACCGCGTCTAACAGTACTAAATATGCGGTCGTTGATTCTGCAACCGGT
AAATACATGGAAGCCACTGTAGTCATTACCGGTACGGCGGCGGCGGTAACCTGTTGGTGCA
GCGGAAGTGGCGGGAGCCGCTACAGCCGATCCGTTAAAAGCACTGGATGCCGCAATCGCT
AAAGTCGACAAATTCCGCTCCTCCCTCGGTGCCGTTCAAACCGTCTGGATTCTGCGGTC
ACCAACCTGAACAACACCACCACCAACCTGTCTGAAGCGCAGTCCCGTATTCAGGACGCC
GACTATGCGACCGAAGTGTTCAACATGTCGAAAGCGCAGATTATCCAGCAGGCGGGCAAC TCCGTGCTGTCTAA

Figure 10

33/96

AACAAAAACCAGTCTGCGCTGTCGACTTCTAT
CGAGCGCCTCTCTTCTGGTCTGCGTATTAACAGCGCTAAAGATGACGCCGCGGGCCAGGC
GATTGCTAACCGCTTCACTTCTAACATCAAAGGTCTGACTCAGGCCGCACGTAACGCCAA
CGACGGTATCTCTCTGGCGCAGACCACTGAAGGCGCGCTGTCTGAAATCAACAACAATT
GCAGCGTGTGCGTGAGTTGACCGTTCAGGCGACGACCGGACTAACTCTGATTCTGACCT
GTCTTCTATTCAAGACGAAATCAAATCCCGTCTGGATGAAATTGATCGCGTTTCCGGTCA
GACCCAGTTCAACGCGGTGAATGTGCTGGCGAAAGATGGTTCGATGAAGATTCAAGTTGG
CGCGAATGATGGGCAGACTATTAGCATTGATTTGCAGAAGATTGACTCTTCTACATTAGG
ACTGAACGGTTTTCTCCGTTTCGGGTCAGTCACTTAACGTTAGTGATTCCATTACTCAAAT
TACCGGTGCCGCCGGGACAAAACCTGTTGGTGTGATTTCACTGCTGTTGCGAAAGATCT
GACTACTGCGACAGGTAAAAACAGTCGATGTTTCTAGCCTGACGTTACACAACACTCTGGA
TGCGAAAGGGGCTGCTACATCACAGTTCGTCGTTCAATCCGGCAATGATTTCTACTCCGC
GTCGATTAATCATAACAGACGGCAAAGTCACGTTGAATAAAGCCGATGTGGAATACACAGA
CACCGATAATGGACTAACGACTGCGGCTACTCAGAAAGATCAACTGATTAAAGTTGCCGC
TGACTCTGACGGCTCGGCTGCGGGATATGTAACATTCCAAGGTAAAACTACGCTACAAC
GGTTTCAACGGCACTTGATGATAATACTGCGGCAAAAGCAACAGATAATAAAGTTGTTGT
TGAATTATCAACAGCAAAACCGACTGCACAGTTCTCAGGGGCTTCTTCTGCTGATCCACT
GGCACTTTTAGACAAAGCTATTGCACAGGTTGATACTTTCCGCTCCTCCCTCGGTGCGGT
GCAAAACCGTCTGGATTCCGCAGTAACCAACCTGAACAACACCACCACCAACCTGTCTGA
AGCGCAGTCCCGTATTCAAGGACGCCGACTATGCTACAGAAGTGTCACATGTGCGAAAGC
GCAGATCATCCAGCAGGCAGGTAACCTCGGTGCTGTCCAAA

Figure 11

34/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGC
TGATCACTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTC
TGTCTTCTGGCTTGCGTATTAAACAGCGCGAAGGATGACGCCGCGGGTCAGGCGATTGCTA
ACCGTTTTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTA
TTTCTGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATTAACAACAACCTACAGCGTA
TTCGTGAACTGACGGTTCAGGCTTCTACCGGGACTAACTCTGATTCCGATCTGGACTCCA
TTCAGGACGAAATCAAATCCCGTCTCGACGAAATTGACCGCGTATCCGGTCAGACCCAGT
TCAACGCGCTGAACGTACTGGCAAAGACGGTTTCGATGAAAATTCAGGTTGGTGCGAATG
ACGGCCAGACTATCACTATTGATCTGAAGAAAATTGACTCTGATACGCTGGGGCTGAATG
GGTTTAAATGTGAACGGCAAAGGGGAAACGGCTAATACGGCAGCAACCCTGAAAGATATGT
CTGGATTACAGCTGCGGCGGCACCAGGGGGAACTGTTGGTGTAACCTCAATATACTGACA
AATCGGCTGTAGCAAGTAGCGTAGATATTCTAAATGCTGTTGCTGGCGCAGATGGAAATA
AAGTTACAACCTAGCGCCGATGTTGGTTTTGGTACACCAGCCGCTGCTGTAACCTATACCT
ACAATAAAGACACTAATTCAATTTCCGCCGCTTCTGATGATATTTCCAGCGCTAACCTGG
CTGCTTTCCCTCAATCCTCAGGCCGAGATACGACTAAAGCTACAGTTACAATTGGTGGCA
AAGATCAAGATGTAAACATCGATAAATCCGGTAATTTAACTGCTGCTGATGATGGCGCAG
TACTTTATATGGATGCTACCGGTAACCTAACTAAAATAATGCTGGTGGTGATACACAAG
CTACTTTGGCTAACTTGCTACTGCTACTGGTGCTAAAGCCGCGACCATCCAACTGATA
AAGGAACATTCAACAGTGACGGTACAGCGTTTGATGGTGATCAATGTCCATTGATACCA
ATACATTTGCAAATGCAGTAAAAATGACACTTATACTGCCACTGTAGGTGCTAAGACTT
ATAGCGTAACAACAGGTTCTGCTGCTGCAGACACCGCTTATATGAGCAATGGGGTTCTCA
GTGATACTCCGCCAACTTACTATGCACAAGCTGATGGAAGTATCACAACCTACTGAGGATG
CGGCTGCCGGTAACTGGTCTACAAAGGTTCCGATGGTAAGTTAACAACGGATACGACTA
GCAAAGCAGAATCAACATCAGATCCGCTGGCAGCTCTTGACGACGCTATCAGCCAGATCG
ACAAATTCCGCTCCTCCCTGGGTGCGGTGCAAACCGTCTGGATTCCGCAGTGACCAACC
TGAACAACACCACTACCAACCTGTCTGAAGCGCAGTCCCGTATTCAGGACGCCGACTATG
CGACCGAAGTGTCCAACATGTGAAAGCGCAGATTATCCAGCAGGCCGGTAACTCCGTGC
TGGCAAAAGCTAACCAGGTTCCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 12

35/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG
 AACCAGTCTG CGCTGTGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC
 GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTA CTCTCTAA CATTAAAGGC
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTCTG TTGCACAGAC CACCGAAGGC
 GCGCTGTCTG AAATCAACAA CAACTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTTCT
 ACCGGAACCTA ACTCTGATTC GGATCTGGAC TCCATTCTAGG ACGAAATCAA ATCCCGTCTT
 GATGAAATTG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT ACTGGCAAAA
 GACGGTTCGA TGAAAATTCA GGTGTTGTCG AATGACGGTG AAATATCAC TATCGACCTG
 AAGAAAATCG ATTCTGATAC TCTGGGTCTG AATGGTTTTA ACGTAAATGG TAAAGGTACT
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTA ACTTCTG CTGGCGCGAA GTTAAACAC
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTA ACTACCG ATGCTGCATT
 CGATAAATTA GGAATGGCG ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAACGC
 TAAATCTGGT GATTTTACTA CCACCAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC
 GCAGGCTACT GATTCAGCTA AAAACGTGA TGCCTTAGCT GCCACCCTTC ATGCTGATGT
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA
 TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAAGCCGC
 GAGCGAAGGT AGTGACGGTG CTTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAAA
 AGCAACTCCT GCGACAACCT CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTA CTCTATCA
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC
 AATTACCTTT AATTCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTTG CCGACTATAC
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGA ACTCCGC
 GGGTAAATC ACTACTGAGA CTACCAGTGC TGGTTCTGCA ACGACCAACC CGCTTGCTGC
 CCTGGACGAC GCTATCAGCT CCATCGACAA ATTCGGTTCT TCCCTGGGTG CTATCCAGAA
 CCGTCTGGAT TCCGAGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CTGAAGCGCA
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTC AACATGTCGA AAGCGCAGAT
 TATCCAGCAG GCCGGTAACT CCGTGCTGGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT
 GTCTCTGCTG CAGGGTTAA

Figure 13

36/96

AACAAATCTCAGTCTTCTCTTAGCTCTGCTA
TTGAGCGTCTGTCTTCTGGTCTGCGTATTAACAGCGCAAAAGACGATGCAGCAGGTCAGG
CGATTGCTAACCGTTTTACGGCAAATATTAAAGGTCTGACCCAGGCTTCCCGTAACGCAA
ATGATGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACC
TGCAGCGTATTCGTGAACCTTCTGTTCAGGCAACTAACGGTACTAACTCTGACAGCGATC
TTTCTTCTATCCAGGCTGAAATTACTCAACGTCTGGAAGAAATTGACCGTGTATCTGAGC
AAACTCAGTTTAACGGCGTGAAAGTCCTTGCTGAAAATAATGAAATGAAAATTCAGGTTG
GTGCTAATGATGGTGAAACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCG
GCCTGGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACAGGCAGTGACCTGATTTCTA
AATTTAAAGCGACAGGTACTGATAATTATGATGTTGGCGGTAAACTTATACCGTGAATG
TGGAGAGCGGCGCGGTTAAGAATGATGCTAATAAAGATGTTTTTGTAAAGCGCAGCTGATG
GATCGCTGACGACCAGTAGTGATACTAAAGTATCCGGTGAAAGTATTGATGCAACAGAAC
TAGCGAAACTTGCAATAAAATTAGCTGACAAAGGCTCCATTGAATACAAGGGCATTACAT
TTACTAACAACACTGGCGCAGAGCTTGATGCTAATGGTAAAGGTGTTTTGACCGCAAATA
TTGATGGTCAAGATGTTCAATTTACTATTGACAGTAATGCACCCACGGGTGCCGGCGCAA
CAATAACTACAGACACAGCTGTTTACAAAAACAGTGCGGGCCAGTTCACCACTACAAAAG
TGGAATAAAGCCGCAACACTCTCTGATCTGGATCTTAATGCAGCCAAGAAAACAGGTA
GCACTTTAGTTGTAAATGGCGCCACCTACAATGTCAGCGCAGATGGTAAAACGGTAACTG
ATACTACTCCTGGTGCCCTAAAGTGATGTATCTGAGCAAATCAGAAGGTGGTAGCCCGA
TTCTGGTAAACGAAGATGCAGCAAAATCGTTGCAATCTACCACCAACCCGCTCGAAACTA
TCGACAAGGCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTACAAAACC
GTTTCGACTCTGCCATCACCAACCTTGGAACACCCGTAAACAACCTGTCTTCTGCCCCGTA
GCCGTATCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCC
TGCAACAAGCGGGTACCTCTGTTCTGGCGCAG

Figure 14

37/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCG
CTGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGT
CTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGCGGGTCAGGCGATTGCT
AACCGTTTTTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGT
ATTTCCGTTGCACAGACCACTGAAGGCGCGCTGTCCGAAATTAACAACAACCTTACAGCGT
ATTCGTGAAGTACGGTTTACGGCTTCTACCGGGACTAATCCGATTTCGGATCTGGACTCC
ATTCAGGACGAAATCAAATCCCGTCTGGACGAAATTGACCGCGTATCCGGCCAGACCCAG
TTCAACGGCGTGAACGTGCTGTCCAAAGATGGCTCGATGAAAATTCAGGTCGGCGCGAAC
GATGGCGAAACGATTACTATTGATCTGAAGAAAATTGACTCTGATACGCTGAATCTGGCT
GGTTTTAACGTTAACGGTAAAGGTTCTGTAGCGAATACAGCTGCGACAAGCGACGATTTA
AAACTGGCTGGTTTTCACTAAGGGCACCACAGATACCAATGGCGTGACCGCGTATACAAAC
ACAATTAGTAATGACAAAGCCAAAGCTTCCGATCTGTTAGCTAATATCACCGATGGATCA
GTGATCACTGGGGGAGGGGCAAACGCTTTTGGCGTGGCTGCAAAGAATGGTTACACCTAT
GATGCAGCAAGTAAATCTTATAGTTTTGCTGCAGATGGTGCCGATTTCAGCGAAGACGTTA
AGCATCATTAATCCAAACACCGGTGATTTCGTGCGCAGGCGACAGTGACTATTGGTGGTAAA
GAGCAGAAAGTTAATATTTCCAGGATGGAAAAATTACTGCGGCAGATGATAATGCGACG
CTGTATTTAGATAAACAGGGAACTTGACAAAACGAAATGCAGGTAACGATACCGCAGCG
ACTTGGGATGGTTTAAATTTCCACAGCGATTCTACCGGTGCGGTTCCAGTTGGGGTTGCA
ACTACAATTACAATTACTTCTGGTACAGCTTCCGGAATGTCTGTTTCAGTCCGCAGGAGCA
GGAATTCAGACCTCAACAAATTTCTCAGATTCTTGCAGGTGGTGCATTTGCGGCTAAGGTA
AGTATTGAGGGAGGCGCTGCTACAGACATTTTGGTAGCAAGTAATGGAAACATAACAGCG
GCTGATGGTAGTGCACTTTATCTTGATGCGACTACTGGTGGATTCACTACAACGGCTGGA
GGAAATACAGCTGCTTCGTTAGATAATTTAATTGCTAACAGTAAGGATGCTACCTTAACC
GTAACCTTCAGGTACCGGCCAGAACACTGTTTATAGCACACAGGAAGTGGCGCTCAGTTC
ACCAAGTTTAGCAAAAGTAGACACAGTCAATGTCACCAACGCACATGTCAGTGCCGAAGGT
ATGGCAAATCTGACAAAAGCAATTTTACCATTGATATGGGCGGTACAGGTACAGTAACT
TACACAGTTTCCAATGGGGATGTGAAAGCTGCTGCAAATGCTGATGTTTATGTCGAAGAT
GGTGCACTTTTCAGCCAATGCTACAAAAGATGTAACTACTTTGAACAAAAAATGGGGCT
ATTACCAACAGCACCGGTGGTACCATCTATGAAACAGCTGATGGTAAGTTAACAACAGAA
GCTACTACTGCATCCAGTTCCACCGCCGATCCCCTGAAAGCTCTGGACGAAGCCATCAGC
TCCATCGACAAATTCGCTCCTCCCTCGGTGCGGTGCAAACCGTCTGGATTCCGCGGTG
ACCAACCTGAACAACACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTTCAGGACGCC
GACTATGCGACCGAAGTGTCCAACATGTGAAAGCGCAGATCATCCAGCAGGCCGGTAAC
TCCGTGCTGGCAAAGCTAACCAAGGTACCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 15

38/96

ATGGCACAAAGTCATTAATACCAACAGC
CTCTCGCTGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATC
GAGCGTCTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGCGGGTCAGGCG
ATTGCTAACCGTTTTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAAC
GACGGTATTTCTGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATTAACAACAACCTTA
CAGCGTGTGCGTGAGCTGACTGTTTCAAGGCGACCACCGGTACTAAGTCTGAGTCTGACCTG
TCTTCTATCCAGGACGAAATCAAATCTCGCCTGGAAGAGATTGATCGTGTTTCAAGTCAG
ACTCAATTTAACGGCGTGAATGTTTTGGCTAAAGATGGGAAAATGAACATTCAAGTTGGG
GCAAATGATGGACAGACTATCACTATTGATCTGAAAAAGATCGATTTCATCTACACTAAAC
CTCTCCAGTTTTGATGCTACAACTTGGGCACCAGTGTTAAAGATGGGGCCACCATCAAT
AAGCAAGTGGCAGTAGGTGCTGGCGACTTTAAAGATAAAGCTTCAGGATCGTTAGGTACC
CTAAAATTAGTTGAGAAAGACGGTAAGTACTATGTAAATGACACTAAAAGTAGTAAGTAC
TACGATGCCGAAGTAGATACTAGTAAGGGTAAAATTAAGTTCAACTCTACAAATGAAAGT
GGAAGTACTCCTACTGCAGCGACGGAAGTAACTACTGTTGGCCGCGATGTAAAATTGGAT
GCTTCTGCACTTAAAGCCAACCAATCGCTTGTCGTGTATAAAGATAAAAGCGGCAATGAT
GCTTATATCATTAGACCAAAGATGTAACTAATCAATCACTTTCAATGCCGCTAAT
ATCAGTGATGCTGGTGTCTTATCTATTGGTGCATCTACAACCGCGCCAAGCAATTTAACA
GCTAACCCGCTTAAGGCTCTTGATGATGCAATTGCATCTGTTGATAAATTCCGCTCTTCT
CTCGGTGCCGTTTCAAGACCGTCTGGATTCTGCCATTGCCAACCTGAACAACACCACTACC
AACCTGTCTGAAGCGCAGTCCCGTATTAGGACGCTGACTATGCGACCGAAGTGTCCAAC
ATGTCGAAAGCGCAGATTATCCAGCAGGCCGGTAACTCCGTGCTGGCAAAGCCAACCAG
GTACCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 16

39/96

AACAAATCTCAGTCTTCTCTGAGCTCCGCCAT
TGAACGTCTCTCTTCTGGCCTGCGTATTAACAGTGCTAAAGATGACGCAGCAGGTGAGGC
GATTGCTAACCGTTTACAGCAAATATTAAAGGTCTGACTCAGGCTTCCCGTAACGCGAA
TGATGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCT
GCAGCGTGACGTGAACTGACTGTTTCAAGCACTAACGGTACTAACTCTGACAGCGATCT
TTCTTCTATCCAGGCTGAAATTACTCAACGTCTGGAAGAAATTGACCGTGTATCTGAGCA
AACTCAGTTTAAACGGCGTGAAAGTCCTTGCTGAAAATAATGAAATGAAATTCAGGTTGG
TGCTAATGATGGTGAAACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCGG
CCTGGACGGTTTTAATATCGATGGCGCGCAGAAAGCAACTGGCAGTGACCTGATTTCTAA
ATTTAAAGCGACAGGTACTGATAACTATGATGTTGGCGGTGATGCTTATACTGTTAACGT
AGATAGCGGAGCTGTTAAAGATACTACAGGGAATGATATTTTGTAGTGACAGCAGATGG
TTCCTGACAACTAAATCTGACACAAACATAGCTGGTACAGGGATTGATGCTACAGCACT
CGCAGCAGCGGCTAAGAATAAAGCACAGAATGATAAATTCACGTTTAAATGGAGTTGAATT
CACAAACAACACTGCAGCGGATGGCAATGGGAATGGTGTATATTCTGCAGAAATTGATGG
TAAGTCAGTGACATTTACTGTGACAGATGCTGACAAAAAAGCTTCTTTGATTACGAGTGA
GACAGTTTACAAAAATAGCGCTGGCCTTTTATACGACAACCAAAGTTGATAACAAGGCTGC
CACACTTTCCGATCTTGATCTCAATGCAGCTAAGAAAACAGGAAGCACGTTAGTTGTTAA
CGGTGCAACTTACGATGTTAGTGCAGATGGTAAACGATAACGGAGACTGCTTCTGGTAA
CAATAAAGTCATGTATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGGTAAACGAAGA
TGCAGCAAAATCGTTGCAATCTACCACCAACCCGCTCGAAACTATCGACAAAGCATTGGC
TAAAGTTGACAATCTGCGTTCTGACCTCGGTGACGTACAAAACCGTTTCGACTCTGCTAT
CACCAACCTTGGCAACACCGTAAACAACCTGTCTTCTGCCCCGTAGCCGTATCGAAGATGC
TGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAACAAGCGGGTAC
CTCTGTTCTGGCGCAG

Figure 17

40/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCA
CTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTT
CTGGCTTGCGTATTAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCTAACCGTT
TCACCTCTAACATTAAAGGCCTGACTCAGGCGGCCCGTAACGCCAACGACGGTATCTCCG
TTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAATTACAGCGTATCCGTG
AACTGACGGTTTACGGCTTCTACCGGGACTAACTCCGATTCGGATCTGGACTCCATTTCAGG
ACGAAATCAAATCCCGTCTGGACGAAATTGACCGCGTATCTGGCCAGACCCAGTTCAACG
GCGTGAACGTACTGGCGAAAGACGGTTCAATGAAAATTTCAGGTTGGTGCGAATGACGGCC
AGACTATCACGATTGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAATGGTTTTTA
ACGTGAATGGTTCCGGTACGATAGCCAATAAAGCGGCGACCATTAGCGACCTGACAGCAG
CGAAAATGGATGCTGCAACTAATACTATACTACAACAAATAATGCGCTGACTGCATCAA
AGGCGCTTGATCAACTGAAAGATGGTGACACTGTTACTATCAAAGCAGATGCTGCTCAAA
CTGCCACGGTTTATACATACAATGCATCAGCTGGTAACTTCTCATTTCAGTAATGTATCGA
ATAATACTTCAGCAAAAGCAGGTGATGTAGCAGCTAGCCTTCTCCCGCCGGCTGGGCAAA
CTGCTAGTGGTGTTTATAAAGCAGCAAGCGGTGAAGTGAACTTTGATGTTGATGCGAATG
GTAAAATCACAATCGGAGGACAGAAAGCATATTTAACTAGTGATGGTAACTTAACTACAA
ACGATGCTGGTGGTGCGACTGCGGCTACGCTTGATGGTTTATTCAAGAAAGCTGGTGATG
GTCAATCAATCGGGTTTAAGAAGACTGCATCAGTCACGATGGGGGGAACAACCTTATAACT
TTAAAACGGGTGCTGATGCTGATGCTGCAACTGCTAACGCAGGGGTATCGTTCACTGATA
CAGCTAGCAAAGAAACCGTTTTTAAATAAAGTGGCTACAGCTAAACAAGGCAAAGCAGTTG
CAGCTGACGGTGATACATCCGCAACAATTACCTATAAATCTGGCGTTCAGACGTATCAGG
CTGTATTTGCCGCAGGTGACGGTACTGCTAGCGCAAAATATGCCGATAAAGCTGACGTTT
CTAATGCAACAGCAACATACACTGATGCTGATGGTGAAATGACTACAATTGGTTTCATACA
CCACGAAGTATTCAATCGATGCTAACAACGGCAAGGTAAGTGTGATTCTGGAAGTGGTA
CGGGTAAATATGCGCCGAAAGTAGGGGCTGAAGTATATGTTAGTGCTAATGGTACTTTAA
CAACAGATGCAACTAGCGAAGGCACAGTAACAAAAGATCCACTGAAAGCTCTGGATGAAG
CTATCAGCTCCATCGACAAATTCGGTTCTTCCCTGGGTGCTATCCAGAACCGTCTGGATT
CCGCAGTCACCAACCTGAACAACACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTC
AGGACGCCGACTATGCGACCGAAGTGTCACATGTGCGAAAGCGCAGATCATTTCAGCAGG
CCGGTAACTCCGTGCTGGCAAAAGCCAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTGC AGGGTTAA

Figure 18

41/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATA
ATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGC
GTATTAACAGCGCGAAGGATGACGCCGAGGTCAGGCGATTGCTAACCGTTTTACTTCTA
ACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCCGTTGCGCAGA
CCACTGAAGGTGCGCTGTCCGAAATCAACAACACTTACAGCGTATTCGTGAGCTGACGG
TTCAGGCTTCTACCGGGACTAACTCCGATTCTGACCTGGACTCCATCCAGGACGAAATCA
AGTCTCGTCTGGACGAAATTGACCGCGTATCCGGTCAGACCCAGTTCAACGGCGTGAACG
TGCTGGCGAAAGACGGTTTCGATGAAAATTAGGTTGGTGCGAATGACGGCCAGACTATCA
CGATTGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAGTGGGTTTAATGTGAATG
GTGGCGGGGCTGTTGCTAACACTGCTGCATCTAAAGCTGACTTGGTAGCTGCTAATGCAA
CTGTGGTAGGCAACAAATATACTGTGAGTGCGGGTTACGATGCTGCTAAAGCGTCTGATT
TGCTGGCTGGAGTTAGTGATGGTGATACTGTTTCAGGCAACCATTAAATAACGGCTTCGGAA
CGGCGGCTAGTGCAACGAATTACAAGTATGACAGTGCAAGTAAGTCTTACTCTTTTGATA
CCACAACGGCTTCAGCTGCCGATGTTTCAGAAATATTTGACCCCGGGCGTTGGTGATACCG
CTAAGGGCACTATTACTATCGATGGTTCTGCACAGGATGTTTCAGATCAGCAGTGATGGTA
AAATTACGTCAAGCAATGGAGATAAACTTTACATTGATACAACTGGGCGCTTAACGAAAA
ACGGCTTTAGTGCTTCTTTGACTGAGGCTAGTCTGTCCACACTTGACGCCAATAATACCA
AAGCGACAACCATTGACATTGGCGGTACCTCTATCTCCTTTACCGGTAATAGTACTACGC
CGAACACTATTACTTATTCAGTAACAGGTGCAAAAGTTGATCAGGCAGCTTTTCGATAAAG
CTGTATCAACCTCTGGAAACGATGTTGATTTCACTACCGCAGGTTATAGCGTCGACGGCG
CAACTGGCGCTGTAAACAAAAGGTGTTGCTCCGGTTTATATTGATAACAACGGGGCGTTGA
CCACATCTGATACTGTAGATTTTTATCTACAGGATGATGGTTTCAGTGACTAACGGCAGCG
GTAAGGCAGTTTATAAAGATGCTGACGGTAAATTGACGACAGATGCTGAAACTAAAGCTG
CAACCACCGCCGATCCCCTGAAAGCTCTGGACGAAGCCATCAGCTCCATCGACAAATTCC
GCTCCTCCCTCGGTGCGGTGCAGAACCGTCTGGATTCCGCGGTACCAACCTGAACAACA
CCACTACCAACCTGTCTGAAGCGCAGTCCCGTATTAGGACGCTGACTATGCGACCGAAG
TATCCAACATGTGAAAGCGCAGATCATCCAGCAGGCCGGTAACTCCGTGCTGGCAAAAG
CTAACCAGGTACCACAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 19

42/96

ATGGCACAAGTCATTAATACCAACAGC
CTCTCGCTGATCACTCAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATC
GAGCGTCTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGAGGTGAGGCG
ATTGCTAACCGTTTTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAAC
GACGGTATTTCTGTTGCACAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACAACCTTA
CAGCGTGTGCGTGAACGACCGTTTCAAGCAACCACCGGTACCAACTCCCAGTCTGACCTG
GACTCTATCCAGGACGAAATTAAATCCCGTCTGGACGAAATTGATCGCGTATCCGGTCAG
ACCCAGTTCAACGGCGTGAACGTGCTGGCAAAAGACGGTTCCATGAAAATTGAGTTGGC
GCGAACGATGGCCAGACCATCACTATCGACCTGAAGAAGATTGACTCTTCTACCTTGAAC
CTGACAGGTTTTTAACGTTAACGGTTCTGGTTCTGTGGCGAATACTGCAGCAACTAAAGCT
GATTTAACCGCTGCTCAACTCTCTGCACCGGGTGCAGCAGACGCAATGGTACAGTTACT
TATACTGTGAGTGTGGTTATAAAGAATCCACTGCTGCAGATGTTATTGCTAGCATCAAA
GACGGCAGTGCTCCGACTTCTGCAATTACTGCAACCATTAAATAATGGCTTCGGTGATTCC
AGTGCGCTGACTTCCAATGACTATACCTATGACCCAGCAAAAGGCGACTTCACTTACGAC
GTAGCTTCAAGCGCCAATAATACTGCTGCCAGGTTGAGTCTTCTCTGACGCCGAAAGCA
GGTGATACCGCAAATCTGAAAGTAACCGTTGGTACGACATCGGTTGATGTCGTTCTGGCC
AGTGATGGTAAGATTACAGCAAAAGATGGTTCTGCATTATATATCGACAGTACAGGTAAC
CTGACTCAGAACAGTGCTGGCTTGACCTCTGCTAAACTGGCTACTCTGACTGGCCTTCAG
GGCTCTGGTGTGCTTCAACCATCACTACTGAAGATGGCACTAATATTGATATTGCTGCT
AACGGTAATATTGGTCTGACCGGTGTTTCGTATCAGTGCTGATTCTCTGCAGTCAGCGACT
AAATCTACGGGCTTTACTGTTGGTACTGGCGCTACAGGTCTGACCGTAGGTACTGATGGT
AAAGTGACTATCGGCGGGACTACTGCTCAGTCCTACACCAGCAAAGATGGTTCCCTGACT
ACTGATAACACCACTAACTGTATCTGCAGAAAGATGGCTCTGTAAACCAACGGTTCAGGT
AAAGCGGTCTATGTAGAAGCGGATGGTGATTTCACTACCGACGCTGCAACCAAAGCCGCA
ACCACCACCGATCCGCTGAAAGCCCTGGATGAGGCAATCAGCCAGATCGATAAGTTCCGT
TCATCCCTGGGTGCTATCCAGAACCGTCTGGATTCCGCGGTACCAACCTGAACAACACC
ACTACCAACCTGTCTGAAGCGCAGTCCCGTATTGAGGACGCCGACTATGCGACCGAAGTG
TCCAACATGTCGAAAGCGCAGATCATTCAGCAGGCCGGTAACTCCGTGCTGGCAAAAGCC
AACCAGGTACCGCAACAGGTTCTGTCTCTGCTGCAGGGCTAA

Figure 20

43/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCAC
TCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTC
TGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGAGGTGAGGCGATTGCTAACCGTTT
TACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCTGT
TGCACAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACAACCTACAGCGTGTGCGTGA
ACTGACCGTTTCAAGCAACCACCGGTACCAACTCCAGTCTGACCTGGACTCTATCCAGGA
CGAAATTAAATCCCGTCTGGACGAAATTGATCGCGTATCCGGTCAGACCCAGTTCAACGG
CGTGAACGTGCTGGCAAAAGACGGTTCCATGAAAATTGAGGTTGGCGCGAACGATGGCCA
GACCATCACTATCGACCTGAAGAAGATTGACTCTTCTACCTTGAACCTGACAGGTTTTAA
CGTTAACGGTTCTGGTTCTGTGGCGAATACTGCAGCAACTAAAGCTGATTTAACCGCTGC
TCAACTCTCTGCACCGGGTGCAGCAGACGCAATGGTACAGTTACTTATACTGTGAGTGC
TGGTTATAAGAATCCACTGCTGCAGATGTTATTGCTAGCATCAAAGACGGCAGTGCTCC
GACTTCTGCAATTACTGCAACCATTAAATGAGCTTCGGTGATTCCAGTGCGCTGACTTC
CAATGACTATACTTATGACCCAGCAAAAGGCGACTTCACCTACGACGTAGCTTCAAGCGC
CAATAATACTGCTGCCCAGGTTTCACTCCTTCCCTGACGCCGAAAGCAGGTGATACCGCAAA
TCTGAAAGTAACCGTTGGTACGACATCGGTTGATGTCGTTCTGGCCAGTGATGGTAAGAT
TACAGCAAAAGATGGTTCTGCATTATATATCGACAGTACAGGTAACCTGACTCAGAACAG
TGCTGGCTTGACCTCTGCTAAACTGGCTACTCTGACTGGCCTTCAGGGCTCTGGTGTTGC
TTCAACCATCACTACTGAAGATGGCACTAATATTGATATTGCTGCTAACGGTAATATTGG
TCTGACCGGTGTTTCGTATCAGTGCTGATTCTCTGCAGTCAGCGACTAAATCTACGGGCTT
TACTGTTGGTACTGGCGCTACAGGTCTGACCGTAGGTACTGATGGTAAAGTGACTATCGG
CGGGACTACTGCTCAGTCCTACACCAGCAAAGATGGTTCCCTGACTACTGATAACACCAC
TAAACTGTATCTGCAGAAAGATGGCTCTGTAACCAACGGTTTCAAGTAAAGCGGTCTATGT
AGAAGCGGATGGTGATTTCACTACCGACGCTGCAACCAAAGCCGCAACCACCACCGATCC
GCTGAAAGCCCTGGATGAGGCAATCAGCCAGATCGATAAGTTCCGTTTATCCCTGGGTGC
TATCCAGAACCGTCTGGATTCCGCGGTACCAACCTGAACAACCACTACCAACCTGTC
TGAAGCGCAGTCCCGTATTGAGGACGCGGACTATGCGACCGAAGTGTCACATGTGCA
AGCGCAGATCATTGAGCAGGCCGGTAACTCCGTGCTGGCAAAAGCCAACAGGTACCGCA
ACAGGTTCTGTCTCTGCTGCAGGGCTAA

Figure 21

44/96

GCGCTGTCGACTTCTATCGAGCGCCTCTCTTCTGGTCTGCGTATTAACAGCGCTAAA
GATGACGCTGCGGGCCAGGCGATTGCTAACCGCTTCACTTCTAACATCAAAGGTCTGACT
CAGGCCGCACGTAAACGCCAACGACGGTATTTCTCTGGCGCAGACGGCTGAAGGCGCGCTG
TCAGAGATTAACAACAACCTGCAGCGTATTCGTGAACTGACCGTTCAGGCCTCTACCGGC
ACGAACTCTGATTCCGACCTGTCTTCTATTTCAGGACGAAATCAAATCCCGTCTTGATGAA
ATTGACCGTGTATCTGGTCAGACCCAGTTCAACGGTGTGAACGTGCTGTGAAAAACGAT
TCGATGAAGATTTCAGATTGGTGCCAATGATAACCAGACGATCAGCATTGGCTTGCAACAA
ATCGACAGTACCACTTTGAATCTGAAAGGATTTACCGTGTCCGGCATGGCGGATTTTCAGC
GCGGCGAAACTGACGGCTGCTGATGGTACAGCAATTGCTGCTGCGGATGTCAAGGATGCT
GGGGGTAAACAAGTCAATTTACTGTCTTACACTGACACCGCGTCTAACAGTACTAAATAT
GCGGTGCTTGATTCTGCAACCGGTAAATACATGGCAGCCACTGTAGTCATTACCAGTACG
GCGGCGGCGGTAACTGTTGGTGCAACGGAAGTGGCGGGAGCCGCTACAGCCGAACCGTTA
AAAGCACTGGATGCCGCAATCGCTAAAGTCGACAAATTCCGCTCCTCCCTCGGTGCCGTT
CAAAACCGTCTGGATTCTGCGGTACCAACCTGAACAACACCACCACCAACCTGTCTGAA
GCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGTCCAACATGTCGAAAGCG
CAGATTATCCAGCAGGCG

Figure 22

45/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATA
ATATCAACAAGAACCAGTCTGCGCTGTGCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGC
GTATTAACAGCGCGAAGGATGACGCCGAGGTGAGGCGATTGCTAACCGTTTACTTCTA
ATATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAATGACGGTATTTCTGTTGCACAGA
CCACTGAAGGCGCGCTGTCCGAAATCAACAACAATTACAGCGTATTCGTGAACTGACGG
TTCAGGCCACTACAGGGACTAATCCGATTCTGACCTGGACTCCATCCAGGACGAAATCA
AATCTCGTCTGGACGAAATTGACCGGTATCCGGTCAGACCCAGTTCAACGGCGTGAACG
TGCTGTCCAAAGATGGTTCAATGAAAATTCAGGTCGGCGCAAATGATGGTGAAACCATCA
CGATTGATCTGAAGAAAATTGACTCTGATACGCTGAATCTGGCTGGTTTTAACGTGAATG
GCGAAGGTGAAACAGCCAATACTGCTGCAACACTTAAAGATATGGTTGGTTTTAAACTCG
ATAATACGGGGGTCACTACAGCTGGAGTTAATAGATATATTGCTGACAAAGCCGTCGCAA
GTAGCACGGATATTTGAATGCGGTAGCTGGTGTGATGGCAGTAAAGTTTCCACGGAGG
CAGATGTTGGTTTTGGTGCAGCTGCCCTGGTACGCCAGTGGAATATACTTATCATAAAG
ATACTAACACATATACGGCTTCTGCTTCAGTTGATGCGACTCAACTGGCGGCATTCCTGA
ATCCTGAAGCGGGTGGTACCACTGCTGCAACAGTAAGTATTGGCAACGGTACAACAGCTC
AAGAGCAAAAAGTCATTATTGCTAAAGATGGTTCTTTAACTGCTGCTGATGACGGTGCCG
CTCTCTATCTTGATGATACTGGTAACTTAAGTAAACTAACGCAGGCACTGATACTCAAG
CTAAACTGTCTGACTTAATGGCAAACAATGCTAATGCCAAAACAGTCATTACAACAGATA
AAGGTACATTTACTGCTAATACGACAAAGTTTGATGGGGTAGATATTTCTGTTGATGCTT
CAACGTTTGCTAACGCCGTTAAAAATGAGACTTACACTGCAACTGTTGGTGTAACTTTAC
CTGCGACATATACAGTCAATAATGGCACTGCTGCATCAGCGTATTTAGTCGATGGAAG
TGAGCAAACTCCTGCCGAGTATTTTGCTCAAGCTGATGGCACTATTACTAGTGGTGAAA
ATGCGGCTACCAGTAAAGCTATCTATGTAAGTGCCAATGGTAACTTAACGACTAATACAA
CTAGTGAATCTGAAGCTACTACCAACCCGCTGGCAGCATTGGATGACGCTATCGCGTCTA
TCGACAAATTCCGTTCTTCCCTGGGTGCTATCCAGAACCGTCTGGATTCCGCAGTCACCA
ACCTGAACAACACCCTACCAACCTGTCTGAAGCGCAGTCCCGTATTACAGACGCCGACT
ATGCGACCGAAGTGTCACATGTGAAAGCGCAGATCATTCAGCAGGCCGTTAACTCCG
TGCTGGCAAAAGCCAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 23

46/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATAATAT
CAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGCGTAT
TAACAGCGCGAAGGATGACGCCGAGGTGAGCGATTGCTAACCGTTTTACTTCTAACAT
TAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCTGTTGCGCAGACCAC
TGAAGGCGCGCTGTCCGAAATTAACAACAACCTTACAGCGTATTCTGTAAGTACGCGTTCA
GGCGACGACCGGAATACTCCACCTCTGACCTGGACTCCATCCAGGACGAAATCAAATC
CCGTCTTGACGAAATTGACCGCGTATCTGGTCAGACCCAGTTCAACGGCGTGAACGTGCT
GTCTAAAGATGGCTCGATGAAAATTCAGGTGCGCGGAACGATGGCGAAACGATTACTAT
TGATCTGAAGAAAATTGACTCTGATACGCTGAATCTGGCTGGTTTTAACGTTAACGGTAA
AGGTTCTGTAGCGAATACCGCTGCGACTACAGATAATCTGACATTGGCTGGTTTTACAGC
GGGTACTAAAGCTGCTGATGGCACCCTAACTTATAGCAAAAATGTCCAGTTTGCCGCCGC
GACTGCAAGCAATGTACTGGCTGCTGCTAAAGATGGCGACGAAATTACGTTGCTGGTAA
TAACGGCACAGGTATAGCTGCAACTGGGGGACTTATACTTATCATAAGGACTCTAACTC
ATACAGCTTTAGCGCAACGGCTGCATCTAAAGATTCTCTGTTGAGCACACTGGCACCAAA
CGCTGGCGATACATTTACCGCTAAAGTGACTATTGGTTCTAAATCGCAAGAAGTTAACGT
TAGCAAGATGGTACGATTACATCCAGCGATGGTAAGGCGCTGTATTTAGATGAGAAGGG
CAACCTGACCCAAACAGGTAGTGGCACAAACAAAGCTGCAACCTGGGATAACCTGATGGC
CAATACAGATACTACAGGCAAAGATGCCTATGGTAACTCTGCGGCAGCAGCTGTTGGGAC
AGTAATCGAAGCAAAAGGAATGACCATCACTTCTGCTGGTGGTAATGCTCAGGTGTTAAA
AGACGCGGCTTATAATGCCGCATATGCGACCTCAATTACTACTGGTACTCCGGGTGATGC
GGGAGCCGCGGGAGCCGCTGCAACTGCGGGTAATGCCGCGGTGGGAGCGCTGGGCGCAAC
GGCAGTTGATAATACCACGGCAGATGTTGCCGATATCTCTATCTCAGCTTCGCAAATGGC
GAGCATCCTTCAGGATAAAGATTTACCTTAAGTGATGGTAGTGATACTTACAACGTGAC
CAGCAATGCTGTCACTATCAATGGCAAAGCAGCAAACATTGATGACAGCGGCGCAATCAC
AGACCAAACCAGTAAAGTTGTCAATTATTTGCTCATACTAACGGTAGCGTGACTAACGA
TACAGGCTCCACTATTTATGCGACAGAAGATGGTAGCCTGACCACCGATGCAGCAACCAA
AGCCGAAACCACCGCCGATCCCTGAAAGCTCTGGACGAAGCCATCAGCTCCATCGACAA
ATTCCGCTCCTCCCTCGGTGCGGTGCAAAACCGTCTGGATTCCGCGGTACCAACCTGAA
CAACACCACCACCAACCTGTCTGAAGCGCAGTCCCGTATTCAGGACGCCGACTATGCGAC
CGAAGTGTCACATGTGAAAGCGCAGATTATCCAGCAGGCCGGTAACCTCCGTGCTGGC
AAAAGCTAACCAGGTACCACAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 24

47/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTG
ATCACTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTG
TCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGCAGGTCAGGCGATTGCTAAC
CGTTTTACTTCTAACATTAAAGGCCTGACTCAGGCGGCCCGTAACGCCAACGACGGTATT
TCTGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATTAACAACAACCTTACAGCGTGTG
CGTGAGCTGACTGTTTCAAGGCGACCACCGGTACCAACTCCCAGTCTGATCTGGACTCTATC
CAGGACGAAATCAAATCCCGTCTGGACGAAATTGACCGCGTATCCGGTCAGACCCAGTTC
AACGGCGTGAACGTGCTGGCAAAGACGGTTCCATGAAAATTCAGGTTGGCGCGAATGAT
GGCCAGACCATCACTATCGACCTGAAGAAGATTGACTCTTCTACGTTGAAACTGACTGGT
TTTAACGTGAATGGTTCTGGTTCTGTGGCGAATACTGCGGCGACTAAAGCGGATTTGGCT
GCTGCTGCAATTGGTACCCCTGGGGCAGCAGATTCTACAGGTGCCATTGCTTACACAGTA
AGTGCTGGGCTGACTAAACTACAGCCGCAGATGTACTGTCTAGCCTCGCTGATGGTACG
ACTATTACAGCCACAGGCGTGAAAAATGGCTTTGCTGCAGGAGCCACTTCCAATGCCTAT
AAACTTAACAAAGATAATAATACATTTACTTATGACACGACTGCTACGACAGCTGAGCTG
CAGTCTTACCTGACTCCGAAAGCGGGCGACACTGCAACATTCAGTGTGAAATTGGTGGT
ACTACACAAGACGTGCTGCTGTCCAGTGATGGCAAACCTCACTGCTAAGGATGGCTCTAAG
CTTTACATTGATACAACCTGGTAATTTAACTCAGAATGGTGGTAATAACGGTGTGGAACA
CTCGCGGAAGCGACTCTGAGTGGTTTAGCTCTGAACAAAAATGGTTTAACGGCTGTTAAA
TCCACAATTACTACAGCTGATAACACTTCGATTGTACTGAATGGTTCAAGCGATGGTACT
GGTAATGCTGGTACTGAAGGTACGATTGCTGTTACAGGCGCTGTAATTAGTTCAGCTGCT
CTGCAATCTGCAAGCAAAACGACTGGTTTCACTGTTGGTACAGTAGACACAGCTGGTTAT
ATCTCTGTAGTACTGATGGGAGTGTTCAGGCATATGATGCTGCGACTTCTGGCAACAAA
GCTTCTTACCAACACTGACGGTACACTGACTACTGATAACACCCTAAACTGTATCTG
CAGAAAGATGGCTCTGTAACCAACGGTTCAGGTAAAGCGGTCTATGTAGAAGCGGATGGT
GATTTCACTACCGACGCTGCAACCAAGCCGCAACCACCACCGATCCGCTGGCCGCTCTG
GATGACGCAATCAGCCAGATCGACAAGTTCGGTTCATCCTTGGGTGCTATCCAGAACCGT
CTGGATTCTGCAGTCACCAACCTGAACAACACCACCACCAACCTGTCTGAAGCGCAGTCC
CGTATTCAGGACGCCGACTATGCGACCGAAGTGTCCAATATGTGAAAGCGCAGATCATC
CAGCAGGCCGGTAACCTCCGTGCTGGCAAAGCCAACCAGGTACCGCAGCAGGTTCTGTCT
CTGCTGCAGGGTTAA

Figure 25

48/96

AACAAATCTCAGTCTTCTCTGAGCTCCGCCATTGAA
CGTCTCTCTTCTGGCCTGCGTATTAAACAGTGCTAAAGATGACGCAGCAGGTCAGGCGATT
GCTAACCGTTTTACAGCAAATATTAAAGGTCTGACTCAGGCTTCCCGTAACGCGAATGAT
GGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCAG
CGTATTCGTGAACTTTCTGTTCAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTCT
TCTATCCAGGCTGAAATTACTCAACGTCTGGAAGAAATTGACCGTGTATCTGAGCAAAT
CAGTTTAAACGGCGTGAAAGTCCTTGCTGAAAATAATGAAATGAAAATTCAGGTTGGTGCT
AATGATGGTGAAACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCGGCCTG
GACGGTTTTTAATATCGATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATTTCTAAATTT
AAAGCGACAGGTACTGATAATTATCAAATTAACGGTACTGATAACTATACTGTTAATGTA
GATAGTGGCGTAGTACAGGATAAAGATGGCAAACAAGTTTATGTGAGTACTGCGGATGGT
TCACTTACGACCAGCAGTGATACTCAATTCAAGATTGATGCAACTAAGCTTGCAAGTGGCT
GCTAAAGATTTAGCTCAAGGGAATAAGATTGTCTACGAAGGTATCGAATTTACAAATACC
GGCACTGTGCTATAGATGCCAAAGGTAATGGTAAATTAACCGCCAATGTTGATGGTAAG
GCTGTTGAATTCATATTTTCGGGGAGTACTGATACATCAGGTACTAGTGCAACCGTTGCC
CCTACGACAGCCCTATACAAAAATAGTGCAGGGCAATTGACTGCAACAAAAGTTGAAAAT
AAAGCAGCGACACTATCTGATCTTGATCTGAACGCTGCCAAGAAAACAGGAAGCACGTTA
GTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAAACGATAACGGAGACTGCT
TCTGGTAACAATAAAGTCATGTATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGGTA
AACGAAGATGCAGCAAAATCGTTGCAATCTACCACCAACCCGCTCGAACTATCGACAAA
GCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCGAC
TCTGCCATCACCAACCTTGGCAACACCGTAAACAACCTGTCTTCTGCCCCGTAGCCGTATC
GAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAACAA
GCGGGTACCTCTGTTCTGGCACAG

Figure 26

49/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATA
ATATCAACAAGAACCAGTCTGCGCTGTGCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGC
GTATTAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCTAACCGTTTTACTTCTA
ACATTAAAGGCCTGACTCAGGCGGCACGTAACGCCAACGACGGTATCTCTCTGGCGCAGA
CCACCGAAGGTGCGCTGTCTGAAATCAACAACAACCTACAGCGGTACGTGAACCTGACCG
TTCAGGCAACCACCGGTACTAACTCCGACTCCGACCTGGCTTCTATTACAGGACGAAATCA
AATCCCGTCTGGATGAAATTGACCGCGTATCTGGTCAGACTCAGTTCAACGGCGTGAACG
TGCTGGCAAAAGACGGTTCCATGAAAATTACAGGTAGGTGCTAACGACGGCCAGACTATCA
CTATTGACCTGAAAAAATCGACTCTGATACTCTGGGCCTGAATGGTTTTAACGTGAATG
GTTCTGGGACGATTACCAACAAAGCAGCAACTGTACGTGATGTTACTCGCGCAGGCGGTA
CATTGGTGAATGGTGCCTATGATATAAAACCACTAACACAGCGCTGACTACAACCTGATG
CCTTCGCGAAATTGAATGATGGTGATGTTGTTACTATCAATAATGGTAAGGATACTGCCT
ATAAATATAATGCTGCTACAGGTGGGTTTACGACGGATGTCTCCATCTCCGGGGATCCTA
CCGCTGCTGACGCTACTGCTAATAAACTGCCCCGTGATGCACTTGCGGCGTCTTTACATG
CTGAGCCGGGTAAACTGTTAATGGTTCTTGGACTACGAATGATGGTACGGTAAATTTG
ATACCGATGCCGATGGTAAGATTTCTATTGGTGGTGTGCTGCTTATGTAGATGCAGCAG
GCAACCTGACCACTAACGCAGCAGGTATGACGACTCAAGCAACAACCTACCGATTTGGTTA
CTGCTGCTGCATCTGCTACTGGTAAGGGTGGATCCCTGACCTTTGGTGACACGACGTATA
AAATTGGTCAGGGTACGGCTGGGGTTGATCCTGATGACGCTTCAGATGATGTACTGGGCA
CCATTTCTTACTCTAAATCAGTAAGCAAGGATGTTGTTCTTGCTGATACTAAAGCAACTG
GTAACACGACAACAGTTGATTTCAACTCCGGTATCATGACTTCAAAGGTTAGTTTCGATG
CAGGTACATCAACTGATACATTCAAAGATGCAGATGGTGCTATCACCAAACTAAAGAAT
ACACCACTTCTTATGCTGTAAATAAAGATACTGGTGAAGTTACCGTTGCTGATTATGCTG
CGGTAGATAGCGCCGATAAGGCTGTTGATGATACTAAATATAAACCGACTATCGGCGCGA
CAGTTAACCTGAATTCTGCAGGTAAATTGACCACTGATACCACAGTGCAGGCACAGCAA
CCAAAGATCCTCTGGCTGCCCTGGACGCTGCTATCAGCTCCATCGACAAATTCGGTTCAT
CCCTGGGTGCTATCCAGAACCCTGCTGGATTCCGCAGTCACCAACCTGAACAACCACTA
CCAACCTGTCCGAAGCGCAGTCCCGTATTACAGGACGCCGACTATGCGACCGAAGTGCCA
ACATGTCGAAAGCGCAGATTATCCAGCAGGCCGGTAACTCCGTGCTGGCAAAAGCCAACC
AGGTACCGCAGCAGGTTCTGTCTCTGCTACAGGGTTAA

Figure 27

50/96

AACAAAAACCAGTCTGCGCTGTCGACTTCTATC
GAGCGCCTTTCTTCTGGTCTGCGTATTAACAGCGCTAAAGATGACGCTGCGGGCCAGGCG
ATTGCTAACCGCTTCACTTCTAACATCAAAGGTCTGACTCAGGCCGCACGTAACGCCAAC
GACGGTATTTCTCTGGCGCAGACCCTGAAGGCGCGCTGTCTGAGATTAACAACAACCTG
CAGCGTGTCGTGAGTTGACTGTACAGGCGACGACCGGGACTAACTCTGATTCTGACCTG
TCTTCTATCCAGGATGAAATCAAATCCCGTTTAAGCGAAATTGACCGTGTATCTGGTCAG
ACTCAGTTTTAACGGCGTGAACGTACTGGCTAAGAATGACACCCTGTCTATTAGGTAGGT
GCAATGACGGTCAGACTATCAATATTGACCTGCAGCAAATCGATTCTCATACACTGGGT
CTGGATGGTTTTAGCGTTAAAAATAATGATGCAGTGAAAACCAAGTGCTGCCGTGAATACT
CTTGGGGGGGGGCGAGGTTCTGTTGCTGTCGACTTCGCAACAACCAAGTTTGACTGCTATC
ACTGGTCTCGGTAGCGGTGCTATCAGCGAAATTGCTAAAGACGATAATGGTGATTACTAC
GCGCATGTACAGGGACTACGGGTAATACTGCTGATGGTTACTATGCTGTGATATCGAC
AAGGCTACCGGTGAGGTGCTCTGAAAGATGGTAACGTAGATACACCGACAGGTACGCCA
ACGACGACAAGCACATATGACTTCACAGACGCTGGTCAAACCGTTTCCTTTGGCACTGAT
GCTGCAACAGCCGGTATCAGCACTGGTGCTTCTCTCGTTAACTTCAGGATGAGAAAGGC
AATGATACTGCTACTTATGCAATCAAAGCACAGATGGCAGCCTGTATGCCGCCAACGTT
GATGAGGCTACCGGTAAAGTCACTGTCAAACCGCCAGCTATACTGATGCTGACGGCAAA
GCAGTGACCGATGCCGCTGTAAACTGGGTGGTGACAATGGCACAACCGAAATTGTTGTC
GATGCTGCGTCAGGTAAACTTACGATGCTGGTGCACTGCAAAACGTTGATCTCTCCAGT
GCAACCAACACGGTAACCGCAATCCCGAACGGTAAACACGCTCTCCGCTGGCTGCCCTT
GACGACGCAATCAGCCAGATCGACAAATTCCGCTCCTCCCTCGGTGCGGTGCAGAACCCT
CTGGATTCGCGGTACCAACCTGAACAACCACTACCAACCTGTCTGAAGCGCAGTCC
CGTATTGAGGACGCTGACTATGCGACCGAAGTATCCAACATGTCGAAAGCGCAGATCATC
CAGCAGGCAGGTAACCTCCGTGCTGTCCAA

Figure 28

51/96

GCGCTGTCGACTTCTATCGAGCGCCTCTCTTCTGGTCTGCGCATTAAACAGCGCTAAAG
ATGACGCTGCGGGCCAAGCGATTGCTAACCGCTTCACTTCTAACATCAAAGGTCTGACTC
AGGCCGCACGTAACGCCAACGACGGTATTTCTCTGGCGCAGACCACTGAAGGCGCACTGT
CTGAAATCAACAACAACCTTGACGCGTTCGTGAACCTGACCGTTCAGGCCACTACCGGTA
CTAACTCTGATTCTGACCTGTCTTCAATACAGGACGAAATCAAATCCCGTCTCGATGAAA
TTGACCGCGTATCCGGTCAGACTCAGTTCAACGGCGTTAATGTTCTTTCCAAAGATGGTT
CAATGAAAATTTCAGGTTGGTGCGAATGATGGTCAAACCTATCTCCATCGATCTGAAGAAAA
TTGATTCTTCAACTTTGGGGCTGAATGGCTTCTCAGTTTCTAAAACTCTCTTAATGTCA
GCAATGCTATCACATCTATCCCGCAAGCCGCTAGCAATGAACCTGTTGATGTTAACTTCG
GTGATACTGATGAGTCTGCAGCAATCGCAGCCAAATTGGGGGTTTCCGATACGTCAAGCC
TGTCGCTGCACAACATCCTTGATAAAGATGGTAAGGCAACAGCTGATTATGTTGTTTCAGT
CAGGTAAAGACTTCTATGCTGCTTCTGTTAATGCCGCTTCAGGTAAAGTAACCTTAAACA
CCATTGATGTTACTTATGATGATTATGCGAACGGTGTGACGATGCCAAGCAAACAGGTC
AGCTGATCAAAGTTTCAGCAGATAAAGACGGCGCAGCTCAAGGTTTTGTACACTTCAAG
GCAAAAACTATTCTGCTGGTGATGCGGCAGACATTCTTAAGAATGGAGCAACAGCTCTTA
AGTTAACTGATCTGAATTTAAGTGATGTTACTGATACTAATGGTAAGGTAACCACAACCTG
CGACTGAGCAATTTGAAGGTGCTTCAACTGAGGATCCGCTGGCGCTTCTGGATAAAGCTA
TTGCATCAGTCGACAAATTCCGGTCTTCTCTAGGTGCCGTGCAGAACCGTCTCGATTCCG
CTATCACCAACCTGAACAACACCACCAACCTGTCTGAAGCGCAGTCCCGTATTTCAGG
ACGCCGACTATGCGACCGAAGTGTCACATGTGCGAAAGCGCAGATCATCCAGCAGGCA

Figure 29

52/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCG
CTGATCACTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGT
CTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGAGGTCAGGCGATTGCT
AACCGTTTTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGT
ATTTCTGTTGCACAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGT
ATTCGTGAAGTACGGTTTCAAGGCGCTACAGGGACTAAGTCCGATTCTGACCTGGACTCC
ATCCAGGACGAAATCAAATCTCGTCTGGACGAAATTGACCGCGTATCTGGTCAGACCCAG
TTCAACGGCGTGAACGTGCTGTCTAAAGATGGCTCGATGAAAATTTCAGGTGCGCGCGAAC
GATGGCGAAACGATTACTATTGATCTGAAGAAAATTGACTCTGATACGCTAAATCTGGCT
GGTTTTAACGTGAATGGTGTCTGGCTCTGTTGATAATGCCAAGGCGACTGGCAAAGATCTT
ACTGATGCTGGTTTTACGGCAAGCGCAGCTGATGCTAATGGCAAATCACTTATACCAA
GACACCGTTACTAAATTCGACAAAGCGACAGCGGCTGATGTATTGGGCAAAGCGGCTGCT
GGCGATAGCATTACCTATGCGGGCACTGATACTGGCTTAGGAGTCGCTGCTGATGCCTCG
ACTTACACCTACAATGCAGCCAATAAGTCTTACACTTTTGATGCTACTGGTGTTGCCAAG
GCGGATGCTGGAACGGCACTGAAAGGGTACTTAGGCGCATCTAACACCGGTAAAATTAAT
ATCGGTGGTACCGAGCAAGAAGTTAACATTGCCAAAGATGGCTCCATCACCGATACCAAT
GGCGATGCGCTGTATCTCGATAGTACCGGCAACTTAACCAAAAATACCGCGAATTTGGGG
GCTGCTGATAAAGCAACTGTAGATAAACTGTTTGCTGGTGCTCAGGATGCAACGATCACC
TTCGATAGCGGCATGACAGCTAAATTCGATCAAACCTGCTGGTACCGTTGATTTCAAAGGC
GCGTCTATTTCTGCTGATGCAATGGCATCAACCTTAAATAATGGTTTCCTATACAGCCAAC
GTAGGTGTAAGGCTTATGCCGTAACCGCTGGCGCAGTTTCAGACAGGTGGCGCAGATGTG
TATAAAGATACCACTGGCGCACTGACGACTGAAGATGACGAAACCGTTACCGCGACCTAC
TACGGTTTTGCTGATGGTAAAGTTTCTGACGGTGAAGGTTCTACTGTCTATAAAGCTGCT
GATGGTTCCATCACTAAAGATGCGACTACCAAGTCTGAAGCAACCACTGACCCCTCTGAAA
GCCCTTGACGACGCAATCAGCCAGATCGACAAATCCGCTCCTCCCTCGGTGCCGTTCAA
AACCGTCTGGATTCCGCCGTACCAACCTGAACAACACCACTACCAACCTGTCTGAAGCG
CAGTCCCCTATTTCAGGACGCCGACTATGCGACCGAAGTGTCCAACATGTGAAAGCGCAG
ATCATTACAGAGGCCGGTAACCTCCGTGCTGGCAAAGCCAACCAGGTACCGCAGCAGGTT
CTGTCTCTGCTGCAGGGTTAA

Figure 30

53/96

AACAAATCTCAGTCTTCTCTTAGCTCTGCTATTGA
GCGTCTCTCTTCTGGCCTGCGTATTAACAGTGCTAAAGATGACGCAGCAGGT CAGGCGAT
TGCTAACCGTTTTACGGCAAATATTAAAGGTCTGACTCAGGCTTCCCGTAACGCGAATGA
TGGTATTTCTGTTGCGCAGACTACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCA
GCGTGACGTGAACTGACTGTT CAGGCAACTAACGGTACTAACTCTGACAGCGATCTTTC
TTCTATT CAGGCAGAAATTACTCAACGTCTGGAAGAAATTGACCGTGATCTGAGCAAAC
TCAGTTTAACGGCGTGAAAGTCCTTGCCGAAAATAATGAAATGAAAATTCAGGTTGGTG
TAATGATGGGGAAACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCGGCCT
GGACGGCTTTAATATCGATGGCGCGCAGAAAGCAACTGGCAGTGACCTGATTTCTAAATT
TAAAGCGACAGGTACTGATAATTATCAAATTAACGGTACTGATACTATACTGTTAATGT
AGATACTGGAGCAGTTCAAAATGAGGATGGTGACGCAATTTTTGTTAGCGCTACCGATGG
TTCTCTGACTACTAAGAGTGATACAAAAGTCGGTGGTACAGGTATTGATGCGACTGGGCT
TGCAAAAGCCGCAGTTTCTTTAGCTAAAGATGCCTCAATTAAATACCAAGGTATTACTTT
CACCAACAAAGGCACTGATGCATTTGATGGCAGTGGTAACGGCACTCTAACCGCTAATAT
TGATGGCAAAGATGTAACCTTTACTATTGATGCGACAGGGAAGGACGCAACATTAAAAAC
GTCTGATCCTGTTTACAAAAATAGTGACAGGTCAGTTCACTACAATAAGGTTGAAAACAA
AGCCGCTACAGCATCGGATCTGGACTTAAATAACGCTAAAAAAGTGGGTAGTTCTTTAGT
TGTAATGGCGCTGATTATGAAGTTAGCGCTGATGGTAAGACAGTAACTGGGCTTGCGAA
AACTATGTATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGGTAAAAGAAGATGCAGC
AAAATCGTTGCAATCTACTACCAACCCGCTCGAAACCATCGACAAGGCATTGGCTAAAGT
TGACAATCTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTTCGACTCTGCTATCACCAA
CCTTGCGAACACCGTAAACAACCTGTCTTCTGCCCGTAGCCGTATCGAAGATGCTGACTA
CGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAACAAGCGGGTACCTCTGT TCTGGCGCAG

Figure 31

54/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAATA
ATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGC
GTATTAACAGCGCGAAGGATGACGCCGAGGTGAGGCGATTGCTAACCGTTTACTTCTA
ACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGATGGTATTTCTGTTGCACAGA
CCACTGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTATCCGTGAACTGACGG
TTCAGGCTTCTACCGGGACTAACTCCGATTCCGATCTGGACTCCATTGAGGACGAAATCA
AATCCCGTCTGGACGAAATTGACCGCGTATCTGGCCAGACCCAGTTCAACGGCGTGAACG
TACTGGCGAAAGACGGTTCAATGAAAATTGAGTTGGTGCGAATGACGGCCAGACTATCA
CGATTGATCTGAAGAAAATTGACTCTGATACGCTGGGGCTGAGTGGGTTTAAATGTGAATG
GTAGCGGGGCTGTGGCTAATACTGCAGCGACTAAATCTGATTTGGCAGCAGCTCAACTCT
TGGCTCCAGGTACTGCTGATGCTAATGGTACAGTTACCTATACTGTTGGCGCAGGCCTGA
AAACATCTACAGCTGCAGATGTAATTGCGAGTTTGGCTAATAACGCAAAAGTTAATGCCA
CAATTGCAAATGGTTTTGGATCGCCAACAGCTACAGATTATACATACAACAGCGCTACAG
GCGATTTTACATATAGTGCAACTATTGCGAGTGGTACAAATTCTGGTGATAGTAACAGTG
CTCAGTTACAATCCTTCCTGACACCAAAAGCGGGCGATACTGCTAACTTAAACGTTAAAA
TTGGTTCTACGTCAATTGACGTTGTATTGGCTAGCGACGGTAAAATTACCGCGAAAGATG
GTTTCAGAACTATTTATTGACGTAGATGGTAACCTCACTCAAAACAATGCTGGGACTGTCA
AAGCAGCCACTCTTGATGCACTGACTAAAACTGGCATAACAACAGGCACACCGAGTGCCG
TATCTACGGTAATTACAACCTGAAGATGAAACAACCTTCACTCTGGCTGGCGGTACTGATG
CTACTACTTCTGGTGCAATCACTGTAGCAAATGCAAGAATGAGTGCTGAGTCTCTTCAAT
CGGCAACTAAGTCCACAGGATTCACAGTTGATGTTGGAGCTACTGGTACCAGCGCAGGCG
ATATTAAAGTTGATAGTAAAGGTATAGTACAACAACACACAGGTACAGGTTTTGAAGACG
CTTACACCAAAGCTGATGGTTCACTGACTACCGATAATACAACCAATCTGTTTTTGCAA
AAGACGGAAGCTGTGACCAATGGTTCAGGTAAAGCAGTCTATGTTTCAGCGGATGGTAATT
TTACTACTGACGCTGAAACTAAAGCTGCAACACCGCCGATCCACTGAAAGCTCTGGACG
AAGCGATCAGCTCCATCGACAAATTCCGTTCTTCCCTCGGTGCGGTGCAAAACCGTCTGG
ATTCCGCAGTCACCAACCTGAACAACACCACTACTAACCTGTCTGAAGCGCAGTCCCGTA
TTCAGGACGCTGACTATGCGACCGAAGTGCCAATATGTCGAAAGCGCAGATCATCCAGC
AGGCCGGTAACTCCGTGCTGGCAAAAGCTAACAGGTACCGCAGCAGGTTCTGTCTCTGC TGCAGGGTTAA

Figure 32

55/96

AACAAAAACCAGTCTGCGCTGTGACTTCTATCGAGCGCCTCTCTT
CTGGTCTGCGCATTAAACAGCGCTAAAGATGACGCTGCGGGCCAGGCGATTGCTAACCGCT
TCACTTCTAACATCAAAGGTCTGACTCAGGCCGCACGTAACGCCAACGACGGTATCTCTC
TGGCGCAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACCTGCAGCGTGTTCTGTG
AGCTGACCGTTTCAGGCCACTACCGGTACTAACTCTGATTCTGACCTGTCTTCAATCCAGG
ACGAAATCAAATCCCGTCTCGATGAAATTGACCGCGTATCCGGTCAGACTCAGTTCAACG
GCGTGAACGTACTGGCAAAAGATAACACCATGAAGATTGAGTTGGTGCGAACGATGGTC
AGACTATATCCATCGACCTGCAAAAAATCGACTCTTCTACTCTTGGTTTGAACGGTTTCT
CCGTTTCTAAAAATGCTCTCGAACTAGCGAAGCGATCACTCAGTTGCCGAACGGTGCGA
ATGCACCAATCGCTGTGAAGATGGATGCGTCTGTTCTGACCGATCTTAACATTACTGATG
CTTCCGCTGTTTTCGCTGCACAACGTAATAAGGTGGTGTGCGAACGTCTACTTATGTTG
TTCAGTATGGCGATAAGAGCTATGCAGCATCTGTTGATGCGGGAGGTACAGTAAAACTGA
ATAAAGCCGACGTAACATATAACGACGCAGCAAATGGTGTTACGAATGCCACCCAGATTG
GTAGTCTGGTTCAGGTTGGTGCTGATGCAACAATGATGCAGTTGGTTTTGTTACCGTGC
AGGGGAAAACTATGTTGCTAATGACTCATTAGTCAATGCTAATGGCGCTGCTGGCGCTG
CAGCAACTAGAGTTACAATTGATGGTGATGGTAGCCTTGGAGCTAACCAGGCTAAAATTG
AACTTAGCCAAAATGGTGCTACTGCTGCAACATCAGAGTTTCGCTGGTGCTTCAACCAACG
ATCCACTGACTCTGCTGGACAAAGCTATCGCATCTGTTGATAAATTCCGTTCTTCTTTGG
GGGCGGTACAGAACCGTCTGAGCTCCGCTGTAACCAACCTGAACAACCACTACCAACC
TGTCTGAAGCGCAGTCCCGTATTGAGGACGCCGACTATGCGACCGAAGTGTCCAACATGT
CGAAAGCGCAGATCATCCAGCAGGCAGGTAACCTCCGTGCTGTCCAAA

Figure 33

56/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAATAATATCAACAAGA
ACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGCCTATTAACAGCG
CGAAGGATGACGCCGAGGTGAGGCGATTGCTAACCGTTTTACTTCTAACATTAAAGGCC
TGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCTGTTGCACAGACCACTGAAGGCG
CGCTGTCCGAAATCAACAACAACCTTACAGCGTATTCGTGAAGTACGGTTCAGGCGACGA
CCGGAACCTAATCCACCTCTGACCTGGACTCCATTACAGGACGAAATCAAATCCCGTCTTG
ATGAAATTGACCGCGTATCCGGCCAAACCCAGTTCAACGGCGTGAACGTACTGTCAAAAG
ATGGCTCGATGAAAATTCAGGTGCGCGCAAATGATGGTGAAACCATCACGATTGATCTGA
AAAAGATCGACTCTTCTACATTGAAGCTGACCAGCTTCAATGTTAACGGTAAAGGCGCTG
TTGATAATGCTAAAGCCACTGAAGCAGATCTGACCGCTGCGGGCTTCTCCCAAGGTGCAG
TCGTCAGTGGCAACAGCACCTGGACTAAATCTACTGTTACTACCTTTAATGCAGCAACAG
CTACCGACGTGCTGGCAAGCGTTAGCGGCGGCAGCACTATTAGCGGTTATACCGGTACAA
ACAATGGATTAGGCGTAGCGGCTTCTACTGCATATACCTACAACGCAACCAGCAAGTCTT
ATTCATTTGACGCAACCGCACTTACCAATGGCGATGGTACTGGGGCCACCCTAAAGTTG
CTGATGTGCTGAAAGCCTATGCAGCAAACGGTGATAATACGGCTCAGATCTCCATCGGCG
GAAGCGCTCAGGACGTTAAAATTGCCAGCGATGGCACCTGACTGACGTCAATGGTGATG
CTTTATATATTGGTTCTGACGGCAACCTGACTAAAAACCAGGCCGGCGGTCCAGATGCGG
CAACGTTGGACGGTATTTTCAACGGTGCGAATGGTAATGCAGCAGTTGATGCGAAGATTA
CATTCGGCAGCGGCATGACCGTTGATTTACCCAGGCTAGCAAAAAGTGGATATTAAGG
GCGCAACGGTATCCGCCGAAGATATGGACACTGCGTTAACTGGGCAGGCTTATACCGTAG
CTAACGGCGCACAGTCTTTTGACGTTGCCGCTGGTGGGGCAGTAACCGCTACTACAGGTG
GCGCTACCGTAAATATTGGTGCTGATGGTGAAGTACGACTGCGACCAACAAGACTGTCA
CAGAACTTATCACGAATTTGCTAACGGCAATATTCTGGATGATGACGGCGCGGCTCTGT
ACAAAGCGGCTGACGTTCTCTGACCACTGAAGCTACTGGTAAATCCGAAGTGACCACGG
ATCCGCTGAAAGCGCTGGACGATGCTATCGCATCCGTAGACAAATCCGCTCCTCCCTCG
GTGCGGTGCAGAACCGTCTGGATTCCGCAGTCACCAACCTGAACAACACCACTACCAACC
TGTCTGAAGCGCAGTCCCGCATTACAGGACGCCGACTATGCGACCGAAGTGTCCAATATGT
CGAAAGCGCAGATCATCCAGCAGGCCGGTAACCTCCGTGCTGGCAAAAGCCAACAGGTAC
CGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 34

57/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCAC
TCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTC
TGGCTTGCGTATTAACAGCGCTAAGGATGACGCCGCGGGTCAGGCGATTGCTAACCGTTT
TACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAAACGCCAACGACGGTATTTCTGT
TGCGCAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTATCCGTGA
ACTGACGGTTTCAGGCTTCTACCGGGACTAACTCCGATTCGGATCTGGACTCCATTTCAGGA
CGAAATCAATCCCGTCTGGACGAAATTGACCGCGTATCTGGCCAGACCCAGTTCAACGG
CGTGAACGTACTGGCGAAAGACGGTTCAATGAAAATTTCAGGTTGGTGCGAATGACGGCCA
GACTATCACTATTGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAGTGGGTTTAA
TGTGAATGGTGGCGGGGCTGTTGCTAATACTGCAGCGACTAAAGATGATTGGTTCGCTGC
ATCAGTTTCAGCTGCGGTAGGTAATGAATACACTGTCTCTGCTGGCCTGTCGAAATCAAC
TGCTGCTGATGTTATTGCTAGTCTCACAGATGGTGCACAGTAACTGCGGCTGGTGTAA
CAATGGTTTTGCTGCAGGGGCAACTGGAGATGCTTATAAATTCAATCAAGCAAACAACAC
TTTTACTTTACAATACCACCTCAACAGCGGCAGAACTCCAATCTTACCTCACGCCTAAGGC
GGGGGATACCGCAACTTTCTCCGTTGAAATTGGTGGCACCAAGCAGGATGTTGTTCTGGC
TAGTGATGGCAAAATCACAGCAAAAGACGGGTCTAAACTTTATATTGACACCACAGGGAA
TTTAACCCAAAACGGTGGAGGTACTTTAGAAGAAGCTACCTCAATGGCTTAGCTTTCAA
CCACTCTGGTCCAGCCGCTGCTGTACAATCTACTATTACTACTGCGGATGGAACCTCAAT
AGTTCTAGCAGGTTCTGGCGACTTTGGAACAACAAAACTGCTGGGGCTATTAATGTCAC
AGGAGCAGTGATCAGTGCTGATGCACCTTCTTCCGCCAGTAAAGCGACTGGGTTTACTTC
TGGCACTTATACCGTAGGTACAGATGGAGTTGTTAAATCTGGTGGCAATGACGTTTATAA
CAAAGCTGACGGGACGGGATTAATACTACTGACAATACCACAAAATATTATTTACAAGATGA
CGGGTCTGTAATAATGGTTCTGGTAAAGCTGTGTATGCTGATGCAACAGGAAAATAAC
TACTGACGCTGAAACTAAAGCCGAAACCACCGCCGATCCCCTGAAAGCTCTGGACGAAGC
GATCAGCTCCATCGACAAAATCCGTTCTTCCCTCGGTGCGGTGCAAAACCGTCTGGATTTC
CGCGGTACCAACCTGAACAACACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTCA
GGACCGCGACTATGCGACCGAAGTGTCCAACATGTGAAAGCGCAGATCATCCAGCAGGC
CGGTAACTCCGTGCTGGCAAAAGCTAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTGCA GGGTTAA

Figure 35

58/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCAC
TCAAAATAATATCAACAAGAACCAGTCTGCGCTGTGCGAGTTCTATCGAGCGTCTGTCTTC
TGGCTTGCCTATTAACAGCGCGAAGGATGACGCCGCGGTCAGGCGATTGCTAACCGTTT
TACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCCGT
TGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTATCCGTGA
ACTGACGGTTCAGGCCACTACCGGTACTAACTCCGATTCTGACCTGGACTCCATCCAGGA
CGAAATCAAATCTCGTCTTGATGAAATTGACCGCGTATCTGGTCAGACCCAGTTCAATGG
CGTGAATGTGTTGTCCAAAGACGGTCAATGAAAATTCAGGTGGGCGCAAATGATGGTGA
AACCATCACGATTGACCTGAAAAAATCGACTCTTCTACACTGAAGCTGACCAGCTTCAA
CGTCAACGGTAAAGGCGCTGTTGATAATGCAAAAGCCACTGAAGCAGATCTGACCGCTGC
GGGCTTCTCCCAAAGTGACGTTGTGCTAGTGGCAATAGCACCTGGACTAAATCTACTGTTAC
TACCTTTAATGCAGCAACAGCTACCGATGTGCTGGCTAGCGTTAGTGGCGGCAGCACTAT
TAGCGGTTATGCTGGCACAACAATGGGTTAGGCGTAGCGGCTTCTACTGCATATACCTA
CAACGCAACCAGCAAGTCTTATTTCATTGACGCAACCGCACTTACTAATGGTGATGGTAC
TGCGGGCTCAACTAAAGTTGCTGATGTTCTGAAAGCCTATGCAGCAAACGCGGATAACAC
GGCTCAGATCTCCATCGGTGGTAGCGCTCAGGAAGTTAAAATTGCCAGCGATGGTACCCT
GACGGATACTAATGGCGATGCTTTATACATTGGTGCTGACGGTAACCTGACGAAAAACCA
GGCCGGCGGGCCAGCCGCGGCAACGTTGGACGGTATTTTCAACGGTGCGAATGGTCATGA
TGCAGTTGATGCGAAGATTACCTTCGGCAGCGGCATGACCGTTGACTTCACCCAGGTTAG
CAACAATGTGGATATTAAGGGCGCGACGGTATCCGCCGAAGATATGAACACTGCGTTAAC
CGGTGAGGCTTATACCGTAGCTAACGGCGCACAGTCTTATGACGTTGCCGCTGATGGTGC
AGTAACTGCTACTACAGGTGGAGCGACCGTAAATATTGGTGCTGAGGGTGAAGTACGAC
TGCGGCCAACAGACTGTACAGAACTTATCACGAATTTGCTAACGGCAATATTCTGGA
TGATGACGGCGCGGCTCTGTATAAAGCGGCTGACGGCTCTCTGACCACTGAAGCTACAGG
TAAATCTGAAGCGACCACGGATCCGCTGAAAGCGCTGGACGATGCTATCGCATCCGTAGA
CAAATTCGTTCTTCCCTGGGTGCCGTGCAGAACCGTCTGGATTCCGCAGTCACCAACCT
GAACAACACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGC
GACCGAAGTGTCCAACATGTGCAAGCGCAGATTATTTCAGCAGGCAGGTAACCTCCGTGCT
GGCAAAAGCTAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 36

59/96

AACAAAAACCAGTCTGCGCTGTCGACTTCTAT
CGAGCGCCTCTCTTCTGGTCTGCGCATTAAACAGCGCTAAAGATGACGCTGCGGGCCAGGC
GATTGCTAACCGCTTCACTTCTAACATCAAAGGTCTGACTCAGGCCGCACGTAACGCCAA
CGACGGTATCTCTCTGGCGCAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACCTT
GCAGCGTGTGCGTGAGTTGACTGTTTCAAGGCGACGACCGGGACTAACTCTGATTCTGACCT
GTCTTCTATTCAAGACGAAATCAAATCCCGTCTGGATGAAATTGACCGTGTTTCCGGTCA
GACCCAGTTCAACGGCGTGAACTGCTGGCTAAAAACGGTTCTATGGCGATTCAAGTTGG
CGCGAATGATGGGCAGACCATCAACATCGACCTGCAGAAAATCGACTCTTCTACTCTGGG
CCTGGGCGGCTTCTCCGTATCTAACATGCACTGAAACTGAGCGATTCTATCACTCAGGT
TGGTGCGAGTGGTTCACTGGCAGATGTGAACTGAGCTCTGTTGCCTCGGCTCTGGGTGT
AGACGCAAGCACTCTGACTCTGCACAACGTACAGACCCAGCTGGCGCAGCAACAGCTAA
CTATGTTGTCTCTTCTGGTTCTGACAATACTCAGTATCTGTTGAAGATAGCTCCGGTAC
AGTTACGCTGAACACCACTGATATAGGTTATACCGATACCGCTAATGGCGTTACTACCGG
TTCCATGACTGGTAAGTACGTTAAAGTTGGAGCTGATGCATTGGGTGCTGCTGTAGGTTA
TGTCACCGTACAGGGACAAACTTCAAAGCTGATGCTGGCGCGCTGGTTAACTCCAAGAA
TGCTGCTGGTAGTCAGAATGTTACTTCTGCAATTGGCGATATTGCTAATAAAGCGAATGC
TAACATTTACACTGGAACCTCTTCTGCAGATCCACTGGCTCTGCTGGACAAAGCTATCGC
ATCTGTTGATAAATTCCGTTCTTCTCTAGGGGCGGTGCAGAACCGTCTGAGCTCTGCTGT
AACCAACCTGAACAACACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTCAAGACGC
CGACTATGCGACCGAAGTGTCCAACATGTCGAAAGCGCAGATCATCCAGCAGGCGGGTAA
CTCCGTGCTGTCTAAA

Figure 37

60/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCA
CTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTT
CTGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGCCGGTCAGGCGATTGCTAACCGTT
TTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAATGACGGTATTTCTG
TTGCACAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTATTCTGTG
AACTGACGGTTTCAGGCTTCTACCGGGACTAACTCTGATTCTGGACTCCATTCAGG
ACGAAATCAAATCCCGTCTCGACGAAATTGACCGCGTATCCGGTCAGACCCAGTTCAACG
GCGTGAACGTACTGGCAAAAGACGGTTCGATGAAAATTTCAGGTTGGTGCGAACGACGGCC
AGACTATCACTATTGATCTGAAGAAAATTGACTCTGATACGCTGGGGCTGAGTGGGTTTA
ACGTAATAGGTAGCGCAGATAAGGCAAGTGTGCGCGCGACAGCTGACGGAATGGTTAAAG
ACGGATATATCAAAGGGTTAACTTCATCTGACGGCAGCACTGCATATACTAAACTACAG
CAAATACTGCAGCAAAAGGATCTGATATTCTTGCGGCGCTTAAGACTGGCGATAAAATTA
CCGCAACAGGTGCAATAGCCTTGCTGATAATGCGACATCGACAACCTTATACTTATAATG
CAACCAGCAATACCTTCTCCTATACGGCTGACGGTGTAACCAAACGAATGCTGCAGCAA
ATCTCATACCTGCAGCAGGGAAAACGACAGCTGCATCAGTTACTATTGGTGGGACAGCAC
AGAATGTAAATATTGATGATTGCGGGCAATATTACTTCAAGTGATGGCGATCAACTTTATC
TGGATTCAACAGGTAACCTGACTAAAAACCAGGCCGGCAACCCGAAAAAAGCAACCGTTT
CTGGGCTTCTCGGAAATACGGATGCGAAAGGTACTGCTGTTAAACAACCATCAAGACAG
AGGCTGGTGTAAACAGTTACAGCTGAAGGTAATACAGGTAAGTGTAAAAATTGAAGGTGCTA
CTGTTTCAGCATCTGCATTTACGGGCATTGCATATTCCGCCAACACCGGTGGGAATACTT
ATGCTGTTGCCGCAAATAATACTACAAATGGTTTCCTGGCGGGGGATGACTTAACCCAGG
ATGCTCAAACCTGTTTCAACCTACTACTCGCAAGCCGATGGCACGGTCACGAATAGCGCAG
GCAAAGAAATCTATAAAGACGCTGATGGTGTCTACAGCACAGAGAATAAAACATCGAAGA
CGTCCGATCCATTGGCTGCGCTTGACGACGCAATCAGCTCCATCGACAAATTCCGTTTCAT
CCTTGGGTGCTATCCAGAACCGTCTGGATTCCGCGGTACCAACCTGAACAACACCACTA
CCAACCTGTCCGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGTTCA
ACATGTGCGAAAGCGCAGATCATCCAGCAGGCCGGTAACCTCCGTGCTGGCAAAAGCTAACC
AGGTACCGCAGCAGTTCTGTCTCTGCTGCAGGGCTAA

Figure 38

61/96

AACAAATCTCAGTCTTCTCTGAGCTCCGCCATTGAACGTCTCTCTTCTGGCCTGCGTA
TTAACAGTGCTAAAGATGACGCAGCAGGTGAGGCGATTGCTAACCGTTTTACAGCAAATA
TTAAAGGTCTGACTCAGGCTTCCCGTAACGCGAATGATGGTATTTCTGTTGCGCAGACCA
CTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCAGCGTGACGTGAACCTGACTGTTT
AGGCAACTAACGGTACTAACTCTGACAGCGATCTTTCTTCTATCCAGGCTGAAATTACTC
AACGTCTGGAAGAAATTGACCGTGTATCTGAGCAAACCTCAGTTTAACGGCGTGAAAGTCC
TTGCTGAAAATAATGAAATGAAAATTCAGGTTGGTGCTAATGATGGTGAAACCATCACTA
TCAATCTGGCAAAAATTGATGCGAAAACCTCTCGGCCTGGACGGTTTTAATATCGATGGCG
CGCAGAAAGCAACTGGCAGTGACCTGATTTCTAAATTTAAAGCGACAGGTACTGATAACT
ATGATGTTGGCGGTGATGCTTATACTGTTAACGTAGATAGCGGAGCTGGGTAATGACTCC
AACTTATTGATAGTGTTTTATGTTTCAAGATAATGCCCAGTACTTTGTCATGCAGCTCCAC
CGATTTTGAGAACGACAGCGACTTCCGTCCCAGCCGTGCCAGGTGCTGCCTCAGATTCAG
GTTATGCCGCTCAATTCGCTGCGTATATCGCTTGCTGATTACGTGCAGCTTTCCCTTCAG
GCGGGATTTCATACAGCGGCCAGCCATCCGTCCATATCACCACGTCAAAGGGTGACAG
CAGGCTCATAAGACGCCCCAGCGTCGCCATAGTGCGTTACCGAATACGTGCGCAACAAC
CGTCTTCCGGAGCCTGTATACGCGTAAAACAGCCAGCGCTGGCGCGATTTAGCCCCGAC
ATAGTCCCCTGTTTCGTCCTTTCCGCGCAGACGATGACGTCACTGCCCGGCTGTATGCG
CGAGGTTACCGACTGCGGCCTGAGTTTTTTAAGTGACGTAAAATCGTGTGAGGCCAACG
CCCATAATGCGGGCAGTTGCCCGGCATCCAACGCCATTTCATGGCCATATCAATGATTTTC
TGGTGCGTACCGGGTTGAGAAGCGGTGTAAGTGAAGTGCAGTTGCCATGTTTTACGGCAG
TGAGAGCAGAGATAGCGCTGATGTCCGGCGGTGCTTTTGCCGTTACGCACCAACCCCGTCA
GTAGCTGAACAGGAGGGACAGCTGATAGAAACAGAAGCCACTGGAGCACCTCAAAAACAC
CATCATACACTAAATCAGTAAGTTGGCAGCATTACCGCGGAGCTGTTAAAGATACTACAG
GGAATGATATTTTTGTTAGTGACAGCAGATGGTTCACTGACAACTAAATCTGACACAAACA
TAGCTGGTACAGGGATTGATGCTACAGCACTCGCAGCAGCGGCTAAGAATAAAGCACAGA
ATGATAAATTCACGTTTAAATGGAGTTGAATTCACAACAACAACCTGCAGCGGATGGCAATG
GGAATGGTGTATATTCTGCAGAAATTGATGGTAAGTCACTGACATTTACTGTGACAGATG
CTGACAAAAAAGCTTCTTTGATTACGAGTGAGACAGTTTACAAAAATAGCGCTGGCCTTT
ATACGACAACCAAAGTTGATAACAAGGCTGCCACACTTTCCGATCTTGATCTCAATGCAG
CTAAGAAAACAGGAAGCACGTTAGTTGTTAACGGTGCAACTTACGATGTTAGTGACAGATG
GTAAAACGATAACGGAGACTGCTTCTGGTAACAATAAAGTCATGTATCTGAGCAAATCAG
AAGGTGGTAGCCCGATTCTGGTAAACGAAGATGCAGCAAAATCGTTGCAATCTACCACCA
ACCCGCTCGAAACTATCGACAAAGCATTGGCTAAAGTTGACAATCTGCGTTCTGACCTCG
GTGCAGTACAAAACCGTTTCGACTCTGCTATCACCACCTTGGCAACACCGTAAACAACC
TGTCTTCTGCCCGTAGCCGTATCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGT
CTCGTGCGCAGATCCTGCAACAAGCGGTACCTCTGTTCTGGCGCAG

Figure 39

62/96

AACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGT
CTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGAGGTCAGGCGATTGCTAACC
GTTTTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTT
CTGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAATTACAGCGTGTGC
GTGAACTGACCGTTTCAGGCAACCACCGGTACCAACTCCCAGTCTGACCTGGACTCTATCC
AGGACGAAATTAAATCCCGTCTGGACGAAATTGACCGCGTATCCGGTCAGACCCAGTTCA
ACGGCGTGAAAGTACTGGCAAAAGACGGTTCCATGAAAATTTCAGGTTGGCGCGAACGATG
GCCAGACCATCACTATCGACCTGAAGAAGATTGACTCTTCTACGCTGAAACTGACTGGTT
TTAACGTGAATGGCAAAGCAGCGGTTGATAATGCTAAAGCGACGGATGCAAATCTGACTA
CCGCCGGTTTTACACAAGGCGTTGTGGATTCAAATGGTAATAGTACTTGGACTAAATCAA
CTACGACTAATTTTCGATGCGGCAACTGCAGTAAACGTACTAGCAGCAGTTAAAGATGGCA
GCACAATCAATTACACCGGTACTGGTAATGGTTTAGGGATTGCTGCAACAAGTGCTTATA
CATATCACGATAGCACTAAATCCTATACCTTTGATTCTACGGGGGCTGCAGTAGCTGGTG
CCGCGTCCAGCCTGCAAGGTACTTTTGGTACAGATACGAATACTGCAAAAATCACCATCG
ATGGTTCTGCTCAAGAAGTAAACATCGCTAAAGATGGGAAAATTACTGATACTGATGGTA
AAGCTTTATATATCGATTCCACTGGTAATTTGACTAAGAACGGCTCTGATACTTTAACTC
AGGCAACATTGAATGATGTCCTTACTGGTGCTAATTCAGTTGATGATACAAGGATTGACT
TCGATAGCGGCATGTCTGTCAACCCTTGATAAAGTGAACAGCACTGTAGATATCACTGGCG
CATCTATTTTCAGCCGCTGCAATGACTAATGAGTTGACAGGTAAGGCCTATACCGTAGTAA
ATGGTGACAGAATCTTACGCTGTAGCTACTAATAACACAGTAAAAACGACTGCTGATGCTA
AAAATGTTTATGTTGATGCTAGTGGTAAATTAATACTGATGACAAAGCCACTGTTACAG
AACTTATCATGAATTTGCGAATGGCAATATCTATGATGATAAAGGCGCTGCTGTTTATG
CGGCGGCGGATGGTTCTCTGACTACAGAACTACAAGTAAATCAGAAGCTACAGCTAACC
CGCTGGCCGCTCTGGACGACGCAATCAGCCAGATCGACAAATTCGGTTCATCCCTGGGTG
CTATCCAGAACCGTCTGGATTCCGCAGTCACCAACCTGAACAACACCACTACCAATCTGT
CTGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGTCCAATATGTCTGA
AAGCGCAGATCATCCAGCAGGCAGGCAACTCCGTGCTGGCAAAA

Figure 40

64/96

AACAAAAACCAGTCTGCGCTGTCGACTTCTATCGAGCGCCTCTCTTCTGGTC
TGCGCATTAACAGCGCTAAAGATGACGCTGCGGGCCAGGCGATTGCTAACCCTTCACTT
CTAACATCAAAGGTCTGACTCAGGCTGCACGTAACGCCAATGACGGTATTTCTCTAGCAC
AGACAGCGGAAGGCGCGCTGTCAGAGATTAACAACAACCTGTCAGCGTGTGCGTGAGTTGA
CCGTGCAGGCAACCACTGGTACCAACTCTGATTCCGATCTCTTCTATTTCAGGATGAAA
TTAAATCTCGTCTGGATGAAATTGACCGCGTCTCTGGTCAGACCCAGTTTAACGGCGTGA
ACGTACTGGCTAAAAACGGTTCTATGGCAATTCAGGTTGGCGCGAACGATGGCCAGACTA
TCTCTATCGACCTGCAGAAAATAGACTCTTCTACTCTGGGTCTGAGCGGCTTCTCTGTTT
CTCAGAACTCCCTGAAACTGAGCGATTCTATCACTACGATCGGCAATACTACTGCTGCAT
CGAAGAACGTGGACCTGAGCGCAGTAGCAACTAACTGGGCGTGAATGCAAGCACCCCTGA
GCCTGCACGAAGTTCAGGACTCTGCTGGTGACGGTACTGGTACCTTCGTTGTTTCTTCTG
GCAGCGACAACCTATGCTGTGTCTGTAGACGCGGCCTCTGGTGCAGTTAACCTGAACACCA
CTGACGTACCTATGATGACGCTACTAATGGTGTTACTGGCGCGACTCAGAACGGTCAGC
TGATCAAAGTAACTTCTGACGCCAACGGTGCAGCTGTTGGTTACGTAACCATTTCAGGGTA
AAAATATCAGGCTGGTGCGACCGGTGTTGACGTTCTGGCGAACAGCGGTGTTGCAGCTC
CAACTACAGCTGTTGATACCGGTACTCTGCAACTGAGCGGTACTGGTGCAACTACTGAGC
TGAAAGGTACTGCAACTCAGAACCCACTGGCACTATTGGACAAAGCTATCGCTTCTGTTG
ATAAATTCGGTTCTTCTCTGGGTGCGGTACAGAATCGTCTGAGCTCTGCTGTAACCAACC
TGAATAACACCACCACTAACCTGTCTGAAGCGCAGTCCCGTATTCAGGATGCCGACTATG
CGACCGAAGTGTCAAATATGTCTAAAGCGCAGATCGTTTCAGCAGGCCGGTAAC

Figure 42

65/96

AACAAATCTCAGTCTTCTCTTAGCTCTGCTATTGAGCGTCTGTCTTCT
GGTCTGCGTATTAACAGCGCAAAAGACGATGCAGCAGGTGAGCGATTGCTAACCGTTTT
ACGGCAAATATTAAAGGTCTGACCCAGGCTTCCCGTAACGCAAATGATGGTATTTCTGTT
GCGCAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCAGCGTATTCGTGAA
CTTTCTGTTTACGGCACTAACGGTACTAAGTCTGACAGCGATCTTTCTTCTATCCAGGCT
GAAATTACTCAACGTCTGGAAGAAATTGACCGTGTATCTGAGCAAACCTCAGTTTAACGGC
GTGAAAGTCCTTGCTGAAAATAATGAAATGAAAATTGAGGTTGGTGCTAATGATGGTGAA
ACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACCTCTCGGCCTGGACGGTTTTAAT
ATCGATGGCGCGCAGAAAGCAACAGGCAGTGACCTGATTTCTAAATTTAAAGCGACAGGT
ACTGATAATTATGATGTTGGCGGTAAAACCTTATACCGTGAATGTGGAGAGCGGCGCGGTT
AAGAATGATGCTAATAAAGATGTTTTTTGTAAGCGCAGCTGATGGATCGCTGACGACCAGT
AGTGATACTAAAGTATCCGGTGAAAGTATTGATGCAACAGAACTAGCGAAACTTGCAATA
AAATTAGCTGACAAAGGCTCCATTGAATACAAGGGCATTACATTTACTAACAACACTGGC
GCAGAGCTTGATGCTAATGGTAAAGGTGTTTTGACCGCAAATATTGATGGTCAAGATGTT
CAATTTACTATTGACAGTAATGCACCCACGGGTGCCGGCGCAACAATAACTACAGACACA
GCTGTTTTACAAAAACAGTGCGGGCCAGTTCACCACTACAAAAGTGGAATAAAGCCGCA
ACACTCTCTGATCTGGATCTTAATGCAGCCAAGAAAACAGGTAGCACTTTAGTTGTAAAT
GGCGCCACCTACAATGTCAGCGCAGATGGTAAAACGGTAACTGATACTACTCCTGGTGCC
CCTAAAGTGATGTATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGGTAAACGAAGAT
GCAGCAAATCGTTGCAATCTACCACCAACCCGCTCGAAACTATCGACAAGGCATTGGCT
AAAGTTGACAATCTGCGTTCTGACCTCGGTGTCAGTACAAAACCGTTTCGACTCTGCCATC
ACCAACCTTGGCAACACCGTAAACAACCTGTCTTCTGCCCCTAGCCGTATCGAAGATGCT
GACTACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAACAAGCGGGTACC
TCTGTTCTGGCGCAG

Figure 43

66/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACT
CAAAATAATATCAACAAGAACCAGTCTGCGCTGTGCGAGTTCTATCGAGCGTCTGTCTTCT
GGCTTGCGTATTAACAGCGCGAAGGATGACGCCGAGGTGAGGCGATTGCTAACCGTTTC
ACCTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCTGTT
GCACAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTATCCGTGAA
CTGACGGTTCAGGCTTCTACCGGGACTAAGTCTGATTTCGGATCTGGACTCCATTACAGGAC
GAAATCAAATCCCGTCTGGACGAAATTGACCGGTATCCGGCCAGACCCAGTTCAACGGC
GTGAACGTGCTGGCGAAAGACGGTTCATGAAAATTCAGGTTGGTGCGAATGACGGCCAG
ACTATCACTATTGATCTGAAGAAAATTGACTCTGATACTCTGGGTTTGAGTGGATTTAAT
GTGAATGGCAAAGGGGCTGTGGCTAACGCAAAAGCGACCGAAGCAGATTTAACGGGGGCT
GGTTTCTCTCAAGGAGCGGTGGATACAAACGGAAATAGTACTTGGACAAAATCAACCACC
ACCAATTACTCAGCTGCAACAACCTGCTGACTTGTATCGACCATTAAGGATGGCTCTACT
GTTACATATGCAGGGACAGACACCGGATTAGGGGTCGCAGCAGCAGGAAATTATACTTAT
GATGCGAACAGTAAATCTTATTCCTTCAATGCCAATGGTCTGACGGGCGCAAATACCGCA
ACTGCACTCAAAGGTTACTTGGGGACAGGTGCTAACACCGCTAAAATTTCTATCGGTGGT
ACAGAGCAGGAAGTGAATATTGCCAAAGATGGCACTATTACAGATACGAATGGTGATGCG
CTCTATCTGGATATTACCGGCAACCTGACTAAGAACTATGCGGGTTCACCACCTGCAGCA
ACGCTGGATAACGTATTAGCTTCCGCAACTGTAAATGCCACTATCAAGTTTGATAGCGGT
ATGACGGTTGATTACACTGCAGGTACTGGCGCGAATATTACAGGTGCATCCATTTCTGCA
GATGACATGGCCGCAAACTGAGCGGAAAGGCGTACACTGTTGCCAATGGTGCTGAGTCT
TATGACGTTGCTGCAGTTACGGGGGCTGTAACAACCTACAGCAGGTAATCACCTGTGTAT
GCCGATGCAGACGGTAAATTAACGACGAGTGCCAGTAATACGGTTACTCAGACTTATCAC
GAGTTTGCTAATGGTAACATTTATGATGACAAAGGCTCGTCACTGTATAAAGCTGCAGAT
GGCTCTCTGACTTCTGAAGCTAAAGGGAAATCTGAAGCAACCGCCGATCCCCTGAAAGCT
CTGGACGAAGCCATCAGCTCCATCGACAAATTCGCTCCTCCCTCGGTGCCGTTCAAAAC
CGTCTGGATTCTGCGGTGACCAACCTGAACAACACCACTACCAACCTGTCTGAAGCGCAG
TCCCGTATTCAGGACGCCGACTATGCGACCGAAGTGTCCAATATGTGCGAAAGCGCAGATC
ATCCAGCAGGCCGGTAACTCCGTGTTGGCAAAAGCTAACAGGTACCGCAGCAGGTTCTG
TCTCTGCTGCAGGGTTAA

Figure 44

67/96

GCGCTGTCGACTTCTATCGAGCGCCTCTCTTCTGGTTTGCGCATTAAACAGCGCTA
AAGATGACGCTGCGGGCCAGGCGATTGCTAACCGCTTCACTTCTAACATCAAAGGTCTGA
CTCAGGCCGCACGTAACGCCAACGACGGTATCTCTCTGGCGCAGACCACTGAAGGCGCAC
TGTCTGAAATCAACAACAACCTTGACGCGTGTTCGTGAACCTGACCGTTTCAGGCCACTACCG
GTACTAACTCTGATTCTGACCTGTCTTCAATCCAGGACGAAATCAAATCCCGCTTGGCTG
AAATCGATCGTGTCTCTGGTCAGACCCAGTTCAACGGCGTGAACGTGCTGGCTAAAAACG
GTTCTCTGAATATTCAGGTTGGCGCGAATGATGGGCAGACCATCTCTATCGATTGTCAGA
AAATAGACTCTTCTGCCCTTGGTTTAAAGTGGTTTTAGTGTGCGCGTGCGGGCGCTAAAAAT
TAAGCGATACAGTGACGCGAGGTGCGCGATGGTTGAGCCGCGCCAGTTAAAGTGGATCTGG
ATGCAGCAGCAACAGATATTGGTACTGCTTTGGGGCAAAGGTTAATGCAAGTTCTTTAA
CGTTGCACAATATCTTAGACAAAGATGGTGCGGCAACTGAGAACTATGTTGTTAGCTATG
GTAGTGATAATTACGCTGCATCTGTTGCAGATGACGGGACTGTAACCTCTTAATAAAACGG
ATATTACTTATTCAGGCGGTGATATTACCGGCGCTACCAAAGATGATACGTTGATTAAAG
TTGCTGCTAATTCTGACGGAGAGGCCGTTGGTTTCGCTACCGTTTCAGGGTAAGAATTATG
AAATTACAGATGGTGTAACCAACAGTCCACTGCTGCACCAACCGATATTGCTCAGACCA
TTGATCTGGATACGGCTGATGAATTTACTGGGGCTTCCACTGCTGATCCACTGGCACTTT
TAGACAAAGCTATTGCACAGGTTGATACTTTCCGCTCCTCCCTCGGTGCCGTTCAAACCC
GTCTGGATTCCGCAGTCACCAACCTGAACAACACTACTACCAACCTGTCTGAAGCGCAGT
CCCGTATTACAGGACGCCGACTATGCGACCGAAGTGTCGAATATGTCGAAAGCGCAGATCA TCCAGCAGGCC

Figure 45

68/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCACT
CAAAATAATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTCT
GGCTTGCGTATTAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCTAACCGTTTT
ACTTCTAATATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAATGACGGTATTTCTCTG
GCGCAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACCTGCAGCGTGTGCGTGAA
CTGACCGTACAGGCGACAACCGGAACGAACCTCCGAATCTGACCTGTCTCTATCCAGGAC
GAAATCAAATCCCGTCTGGAAGAGATTGACCGCGTATCCGGCCAGACTCAGTTCAACGGC
GTGAATGTGCTGGCAAAGACGGCACCATGAAAATTGAGGTAGGCGCGAACGATGGTCAG
ACTATCTCTATCGATCTGAAAAAATCGACTCTTCAACCCTGGGCCTGACCGGTTTTGAT
GTTTCGACGAAAGCGAATATTTCTACGACAGCAGTAACGGGGGCGGCAACGACCACTTAT
GCTGATAGCGCCGTTGCAATTGATATCGGAACGGATATTAGCGGTATTGCTGCTGATGCT
GCGTTAGGAACGATCAATTTGATAATACAACAGGCAAGTACTACGCACAGATTACCAGT
GCGGCCAATCCGGGCCTTGATGGTGCTTATGAAATCCATGTTAATGACGCGGATGGTTCC
TTCAGTGTAGCAGCGAGTGATAAACAAGCGGGTGCTGCTCCGGGTACTGCTCTGACAAGC
GGTAAAGTTTCAGACTGCAACCACCACGCCAGGTACGGCTGTTGATGTCACTGCGGCTAAA
ACTGCTCTGGCTGCAGCAGGTGCTGACACGAGTGGCCTGAAACTGGTTCAACTGTCCAAC
ACGGATTCCGCAGGTAAAGTGACCAACGTGGGTACGGCCTGCAGAATGACAGCGGCACT
ATCTTTGCAACCGACTACGATGGCACCCTGTGACCACGCCGGGCGCAGAGACTGTGACT
TACAAAGATGCTTCCGGTAACAGCACCCTGCGGCTGTCACTGGGTGGCTCTGATGGC
AAAACCAATCTGGTTACCGCCGCTGACGGCAAACGTACGGTGCGACTGCACTGAATGGT
GCTGATCTGTCCGATCCTAATAACACCGTTAAATCTGTTGCAGACAACGCTAAACCGTTG
GCTGCCCTGGATGATGCAATTGCGATGGTCGACAAATTCCGCTCCTCCCTCGGTGCGGTG
CAAAACCGTCTGGATTCCGCAGTCACCAACCTGAACAACCACTACCAACCTGTCTGAA
GCGCAGTCCCGTATTGAGGACGCCGACTATGCGACCGAAGTGTCCAACATGTGAAAGCG
CAGATTATCCAGCAGGCGAGTAACCTCCGTGCTGTCCAAAGCTAACAGGTTCCGCAGCAG
GTTCTGTCTCTGCTGCAGGGTTAA

Figure 46

69/96

AACAAAAACCAGTCTGCGCTGTCGACTTCTATCGAGCGCCTCTCTTCTGGT
CTGCGTATTAACAGCGCTAAAGATGACGCCGCGGGCCAGGCGATTGCTAACCGCTTTACT
TCTAACATCAAAGGTCTGACTCAGGCCGCACGTAACGCCAACGACGGTATTTCTCTGGCG
CAGACGGCTGAAGGCGCGCTGTCAGAGATTAACAACAACCTTGCAGCGTATTCGTGAACTG
ACCGTTTCAGGCCTCTACCGGCACGAACCTCTGATTCCGACCTGTCTTCTATTTCAGGACGAA
ATCAAATCCCGTCTTGATGAAATTGACCGTGTATCTGGTCAGACCCAGTTCAACGGTGTG
AACGTGCTGTGCAAAAAACGATTTCGATGAAGATTCAGATTGGTGCCAATGATAACCAGACG
ATCAGCATTGGCTTGCAACAAATCGACAGTACCACTTTGAATCTGAAAGGATTTACCGTG
TCCGGCATGGCGGATTTTCAGCGCGGCGAAACTGACGGCTGCTGATGGTACAGCAATTGCT
GCTGCGGATGTCAAGGATGCTGGGGGTAAACAAGTCAATTTACTGTCTTACACTGACACC
GCGTCTAACAGTACTAAATATGCGGTCGTTGATTCTGCAACCGGTAAATACATGGAAGCC
ACTGTAGCCATTACCGGTACGGCGGCGGCGGTAAGTGTGGTGCAGCGGAAGTGGCGGGA
GCCGCTACAGCCGATCCGTTAAAAGCACTGGATGCCGCAATCGCTAAAGTCGACAAATTC
CGCTCCTCCCTCGGTGCCGTTCAAACCGTCTGGATTCTGCGGTACCAACCTGAACAAC
ACCACCACCAACCTGTCTGAAGCGCAGTCCCGTATTCAGGACGCCGACTATGCGACCGAA
GTGTCCAACATGTCGAAAGCGCAGATTATCCAGCAGGCCGGTAAGTCCGTGCTGGCAA

Figure 47

70/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCACTC
AAAATAATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTCTG
GCTTGCGTATTAAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCTAACCGTTTTA
CCTCTAACATTAAAGGTCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCTGTTG
CACAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTATCCGTGAAC
TGACGGTTTCAGGCTTCTACCGGGACTAACTCCGATTCGGATCTGGACTCCATTTCAGGACG
AAATCAAATCCCGTCTGGACGAAATTGACCGCGTATCCGGTCAAACCCAGTTCAACGGTG
TGAACGTACTGGCGAAAGACGGTTCGATGAAAATTTCAGGTTGGTGCGAATGACGGCCAGA
CTATCACGATTGATCTGAAGAAAATTGACTCAGATACGCTGGGGCTGAATGGTTTCAACG
TTAATGGCAAAGGCACTATTGCGAACAAAGCTGCTACAGTCAGCGATCTGACCGCTGCTG
GTGCAACGGGAACAGGTCCTTATGCTGTGACCACAAACAATACAGCACTCAGCGCTAGCG
ATGCACTGTCTCGCCTGAAAACCGGAGATACAGTTACTACTACTGGCTCGAGTGCTGCGA
TCTATACTTATGATGCGGCTAAAGGGAACCTTCACCACTCAAGCAACAGTTGCAGATGGCG
ATGTTGTTAACTTTTCGAATACTCTGAAACCAGCGGCTGGCACTACTGCATCAGGTGTTT
ATACTCGTAGTACTGGTGATGTGAAGTTTGATGTAGATGCTAATGGCGATGTGACCATCG
GTGGTAAAGCCGCGTACCTGGACGCCACTGGTAACCTATCTACAAACAACCCCGGCATTG
CATCTTCAGCGAAATTGTCCGATCTGTTTGCTAGCGGTAGTACCTTAGCGACAACCTGGTT
CTATCCAGCTGTCTGGCACAACCTTATAACTTTGGTGCAGCGGCAACTTCTGGCGTAACCT
ACACCAAACTGTAAGCGCTGATACTGTACTGAGCACAGTGCAGAGTGCTGCAACGGCTA
ACACAGCAGTTACTGGTGCGACAATTAAGTATAATACAGGTATTTCAGTCTGCAACGGCGT
CCTTCGGTGGTGTAATACTAATGGTGCTGGTAATTCGAATGACACCTATACTGATGCAG
ACAAAGAGCTCACCACAACCGCATCTTACACTATCAACTACAACGTCGATAAGGATACCG
GTACAGTAACTGTAGCTTCAAATGGCGCAGGTGCAACTGGTAAATTTGCAGCTACTGTTG
GGGCACAGGCTTATGTTAACTCTACAGGCAAACTGACCACTGAAACCACCAGTGCAGGCA
CTGCAACCAAGATCCTCTGGCTGCCCTGGATGAAGCTATCAGCTCCATCGACAAATTCC
GTTTCATCCCTGGGTGCTATCCAGAACCGTCTGGATTCCGCGGTTACCAACCTGAACAACA
CCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACCGAAG
TGTCCAACATGTGAAAGCGCAGATTATCCAGCAGGCGGTAACCTCCGTGCTGGCAAAAG
CCAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 48

71/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCAC
TCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTC
TGGCTTGCGTATTAAACAGCGCGAAGGATGACGCCGAGGTGAGGCGATTGCTAACCGTTT
TACTTCTAATATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAATGACGGTATTTCTGT
TGCACAGACCACTGAAGGCGCGTGTCCGAAATCAACAACAACCTTACAGCGTGTGCGTGA
ACTGACCGTTTCAGGCGACCAACCGGTACCAACTCCCAGTCTGATCTGGACTCTATCCAGGA
CGAAATCAATCCCGTCTGGACGAAATTGACCGCGTATCCGGTCAGACTCAGTTCAACGG
CGTGAACGTACTGGCAAAAGACGGTTCATGAAAATTCAGGTTGGCGCGAATGATGGCCA
GACCATCACTATCGACCTGAAGAAGATTGACTCTTCTACGTTGAACTGACTGGTTTTAA
CGTGAATGGTTCTGGTTCTGTGGCGAATACTGCGGCGACTAAAGACGAACTGGCTGTGC
TGCTGCGGCGGCGGGTACAACCTCTGCTGTGCGTACTGACGGCGTGACCAATATACCGT
AGACGCAGGGCTTAACAAAGCCACAGCAGCAAACGTGTTTGCAAACCTTGACAGTGGTGC
TGTTGTTGATGCTAGCATTTCCAACGGTTTTGGTGACGAGCAGCCACAGACTACACCTA
CAATAAAGCTACAAATGATTTCACTTTCAATGCCAGCATTGCTGCTGGTGCTGCGGCCGG
TGATAGTAACAGCGCAGCTCTGCAATCCTTCCTGACTCCAAAAGCAGGTGATACAGCTAA
CCTGAGCGTCAAAATCGGTACGACATCTGTTAATGTTGTTCTGGCGAGCGATGGCAAAAT
TACAGCGAAAGATGGCTCAGCTCTGTATATCGACTCAACGGGTAACCTGACTCAGAACAG
CGCAGGCACTGTAACAGCAGCAACCCCTGGATGGACTGACCAAAAACCATGATGCGACAGG
AGCTGTTGGTGTGATATCACGACCGCAGATGGCGCAACTATCTCTGCGCAGGCTCTGC
TAACGCGGCAACAGGTACTCAATCAGGTGCAATTACACTGAAAAATGTTGATATCAGTGC
TGATGCTCTGCAGTCTGCTGCGAAAGGTACTGTTATCAATGTTGATAATGGTGCTGATGA
TATTTCTGTTAGTAAAACCGGGTGTGCTTACTACCGGAGGTGCGCCTACTTATACTGATG
CTGATGGTAAATTAACGACAACCAACACCGTTGATTATTTCTGCAAACCTGATGGCAGCG
TAACCAATGGTTCTGGTAAAGGGGTTTACACCGATGCAGCTGGTAAATTCATAACGACG
CTGCAACCAAAGCCGCAACCAACCGGATCCGCTGAAAGCCCTTGATGACGCAATCAGCC
AGATCGATAAGTTCCGTTTCATCCCTGGGTGCTATCCAGAACCGTCTGGATTCCGCGGTTA
CCAACCTGAACAACCACTACCAACCTGTCCGAAGCGCAGTCCCGTATTCAGGACGCCG
ACTATGCGACCGAAGTGTCCAATATGTGCAAGCGCAGATCATCCAGCAGGCCGGTAAC
CCGTGTTGGCAAAAGCTAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 49

72/96

AACAAATCTCAGTCTTCTCTTAGCTCTGCTATTGAGCGTCTGTCTTCTGGT
CTGCGTATTAAACAGCGCAAAAGACGATGCAGCAGGTGAGGCGATTGCTAACCGTTTTACG
GCAAATATTAAAGGTCTGACCCAGGCTTCCCGTAACGCGAATGATGGTATTTCTGTTGCG
CAGACCACTGAAGGTGCGCTGAATGAAATTAACAACAACCTGCAGCGTATTCGTGAACCT
TCTGTTTACAGCAACTAACGGTACTAACTCTGACAGCGATCTTTCTTCTATCCAGGCTGAA
ATTACTCAACGTCTGGAAGAAATTGACCGTGTATCTGAGCAAACCTCAGTTTAACGGCGTG
AAAGTCCTTGCTGAAAATAATGAAATGAAAATTCAGGTTGGTGCTAATGATGGTGAAACC
ATCACTATCAATCTGGCAAAAATTGATGCGAAAACTCTCGGCCTGGACGGTTTTAATATC
GATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATTTCTAAATTTAAAGCGACAGGTACT
GATAATTATCAAATTAACGGTACTGATAACTATACTGTTAATGTAGATAGTGGAGTAGTA
CAGGATAAAGATGGCAAAACAAGTTTTATGTGAGTGCTGCGGATGGTTCACTTACGACCAGC
AGTGATACTCAATTCAAGATTGATGCAACTAAGCTTGCACTGGCTGCTAAAGATTTAGCT
CAAGGTAATAAGATTGTCTACGAAGGTATCGAATTTACAAATACCGGCACTGGCGCTATA
CCTGCCACAGGTAATGGTGAATTAACCGCCAATGTTGATGGTAAGGCTGTTGAATCACT
ATTTCCGGGAGTGCTGATACATCAGGTACTAGTGCAACCGTTGCCCCCTACGACAGCCCTA
TACAAAAATAGTGCAAGGCAATTGACTGCAACAAAAGTTGAAAATAAAGCAGCGACACTA
TCTGATCTTGATCTGAACGCTGCCAAGAAAACAGGAAGCACGTTAGTTGTTAACGGTGCA
ACTTACGATGTTAGTGCAAGTGGTAAAACGATAACGGAGACTGCTTCTGGTAACAATAAA
GTCATGTATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGGTAACGAAGATGCAGCA
AAATCGTTGCAATCTACCACCAACCCGCTCGAAACTATCGACAAAGCATTGGCTAAAGTT
GACAATCTGCGTTCTGACCTCGGTGCAAGTACAAAACCGTTTCGACTCTGCCATCACCAC
CTTGGCAACACCGTAACAACCTGTCTTCTGCCCGTAGCCGTATCGAAGATGCTGACTAC
GCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCCTGCAACAAGCGGTACCTCTGTT CTGGCACAG

Figure 50

73/96

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCAC
TCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCGAGCGTCTGTCTTC
TGGCTTGCGTATTAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCTAACCGTTT
CACCTCTAACATTAAAGGCTGACTCAGGCGGCCCGTAACGCCAACGACGGTATCTCCGT
TGCGCAGACCACCGAAGGCGCGCTGTCCGAAATCAACAACAACCTTACAGCGTGTGCGTGA
ACTGACGGTACAGGCCACTACCGGTACTAACTCTGAGTCTGATCTGTCTTCTATCCAGGA
CGAAATTAAATCCCGTCTGGATGAAATTGACCGCGTATCTGGTCAGACCCAGTTCAACGG
CGTGAACGTGCTGGCAAAAAATGGCTCCATGAAAATCCAGGTTGGCGCAAATGATAACCA
GACTATCACTATCGATCTGAAGCAGATTGATGCTAAAACCTCTTGGCCTTGATGGTTTTAG
CGTTAAAAATAACGATACAGTTACCACTAGTGCTCCAGTAACTGCTTTTGGTGCTACCAC
CACAAACAATATTAACTTACTGGAATTACCCTTTCTACGGAAGCAGCCACTGATACTGG
CGGAACTAACCCAGCTTCAATTGAGGGTGTTTATACTGATAATGGTAATGATTACTATGC
GAAAATCACCGGTGGTGATAACGATGGGAAGTATTACGCAGTAACAGTTGCTAATGATGG
TACAGTGACAATGGCGACTGGAGCAACGGCAAATGCAACTGTAACCTGATGCAAATACTAC
TAAAGCTACAACCTATCACTTCAGGCGGTACACCTGTTTACAGTTGATAATACTGCAGGTTT
CGCAACTGCCAACCTTGGTGCTGTTAGCTTAGTAAACCTGCAGGATTCCAAGGGTAATGA
TACCGATACATATGCGCTTAAAGATACAAATGGCAATCTTTACGCTGCGGATGTGAATGA
AACTACTGGTGCTGTTTCTGTTAAACCTATTACCTATACTGACTCTTCCGGTGCCGCCAG
TTCTCCAACCGCGGTCAAACCTGGGCGGAGATGATGGCAAAACAGAAGTGGTCGATATTGA
TGGTAAACATACGATTCTGCCGATTTAAATGGCGGTAATCTGCAAACAGGTTTGACTGC
TGGTGGTGAGGCTCTGACTGCTGTTGCAAATGGTAAAACCACGGATCCGCTGAAAGCGCT
GGACGATGCTATCGCATCTGTAGACAAATCCGTTCTTCCCTCGGTGCGGTGCAAAACCG
TCTGGATTCCGCGGTTACCAACCTGAACAACCACTACCAACCTGTCTGAAGCGCAGTC
CCGTATTACAGGACGCCGACTATGCGACCGAAGTGTCCAATATGTCGAAAGCGCAGATCAT
CCAGCAGGCCCGGTAACCTCCGTGTTGGCAAAAGCTAACCAGGTACCGCAGCAGGTTCTGTCT
TCTGCTGCAGGGTTAA

Figure 51

74/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACT
CAAAATAATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTCT
GGCTTGCGTATTAAACAGCGCAAGGATGACGCCGAGGTCAGGCGATTGCTAACCGTTTT
ACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTCTGTT
GCGCAGACCACCGAAGGCGCGCTGTCTGAAATCAACAACAACCTACAGCGTATTCTGTGAA
CTGACGGTTTCAGGCTTCTACCGGGACTAACTCTGATTCTGGACTCCATTTCAGGAC
GAAATCAAATCCCGTCTGGACGAAATTGACCGCGTATCCGGTCAAACCCAGTTCAACGGT
GTGAACGTACTGGCGAAAGACGGTTTCGATGAAAATTTCAGGTTGGTGCGAATGACGGCCAG
ACTATCACTATTGATCTGAAGAAAATTGACTCTGATACGCTGGGGCTGAATGGTTTTAAC
GTTAACGGCAAAGGTACTATTGCGAACAAGCGGCAACCATTAGTGATCTGGCGGCGACG
GGGGCGAATGTTACTAACTCAAGCAATATTGTTGTACGACAAAGTTCAATGCCTTGGAT
GCAGCGACTGCATTTAGCAAACCTCAAAGATGGTGATTCTGTTGCCGTTGCTGCTCAGAAA
TATACTTATAACGCATCGACCAATGATTTTACGACAGAAAATACAGTAGCGACAGGCCT
GCAACGACAGATCTTGGCGCTACTCTGAAGGCTGTGCTGGGCAGAGTCAATCAGGTACA
TATACCTTTGCAAATGGTAAAGTTAACTTTGATGTTGATGCAAGCGGTAATCACTATT
GGCGGCGAAAAGGCTTTCTTGTTGGTGGAGCGCTGACTACTAACGATCCCACCGGCTCC
ACTCCAGCAACGATGTCTTCCCTGTTTAAGGCCGCGGATGACAAAGATGCCGCTCAATCC
TCGATTGATTTTGGCGGGAAAAAATACGAATTTGCTGGTGGCAATTCTACTAATGGTGGC
GGCGTTAAATTCAAAGACACGGTGTCTTCTGACGCGCTTTTGGCTCAGGTTAAAGCGGAT
AGTACTGCTAATAATGTAAAAATCACCTTTAACAATGGTCCTCTGTCACTCACTGCATCG
TTCCAAAATGGTGTATCTGGCTCCGCGGCATCGAATGCAGCCTACATTGATAGCGAAGGC
GAACTGACAACTACTGAATCCTACAACACAAATTATCCGTAGACAAAGACACGGGGGCT
GTAAGTGTTACAGGGGGGAGCGGTACGGGTAAATACGCCGCAAACGTGGGTGCTCAGGCT
TATGTAGGTGCAGATGGTAAATTAACCACGAATACTACTAGTACCGGCTCTGCAACCAAA
GATCCACTAAATGCGCTGGATGAGGCAATTGCATCCATCGACAAATTCCGTTCTTCCCTG
GGGGCTATCCAGAACCGTCTGGATTCCGCAGTCACCAACCTGAACAACACCACTACCAAC
CTGTCTGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGTCACATG
TCGAAAGCGCAGATCATCCAGCAGGCCGGTAACTCCGTGTTGGCAAAAGCTAACCAGGTA
CCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 52

75/96

AACAAGAACCAGTCTGCGCTGTCTGAGTTCTATCGAGCGTCTGTC
TTCTGGCTTGCGTATTAAACAGCGCGAAGGATGACGCCGCGGGTCAGGCGATTGCTAACCG
TTTTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACGACGGTATTTC
TGTTGCGCAGACCACCGAAGGCGCGCTGTCCGAAATTAACAACAACCTTACAGCGTGTGCG
TGAGCTGACTGTTTCAGGCGACCACCGTACTAAGTCTGAGTCTGACCTGTCTTCTATCCA
GGACGAAATCAATCTCGCCTGGAAGAGATTGATCGTGTTCAGTCTGACTCAATTTAA
CGGCGTGAATGTTTTGGCTAAAGATGGGAAAATGAACATTGAGGTTGGGGCAAGTGATGG
ACAGACTATCACTATTGATCTGAAAAAGATCGATTCTACACTAAACCTCTCCAGTTT
TGATGCTACAACTTGGGCACCAAGTGTAAAGATGGGGCCACCATCAATAAGCAAGTGGC
AGTAGATGCTGGCGACTTTAAAGATAAAGCTTCAGGATCGTTAGGTACCCTAAAATTAGT
TGAGAAAGACGGTAAGTACTATGTAAATGACACTAAAAGTAGTAAGTACTACGATGCCGA
AGTAGATACTAGTAAGGGTGAAATTAACCTCAACTCTACAAATGAAAGTGGAAGTACTCC
TACTGCAGCGACGGAAGTAACTACTGTTGGCCGCGATGTAAATTTGGATGCTTCTGCACT
TAAAGCCAACCAATCGCTTGTCGTGTATAAAGATAAAAGCGGCAATGATGCTTATATCAT
TCAGACCAAAGATGTAACAATAATCAATCAACTTTCAATGCCGCTAATATCAGTGATGC
TGGTGTTTTATCTATTGGTGATCTACAACCGCGCCAAGCAATTTAACAGCTGACCCGCT
TAAGGCTCTTGATGATGCAATTGCATCTGTTGATAAATTCGCTCTTCTCTCGGTGCCGT
TCAGAACCGTCTGGATTCTGCCATTGCCAACCTGAACAACCACTACCAACCTGTCTGA
AGCGCAGTCCCGTATTGAGGACGCTGACTATGCGACCGAAGTGTCCACATGTGCGAAAGC
GCAGATTATCCAGCAGGCCGGTAAGTCCGTGCTGGCAAAA

Figure 53

ATGGCACAAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAAA
TAATATCAACAAGAACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTCTGGCTT
GCGTATTAACAGCGCGAAGGATGACGCAGCGGGTCAGGCGATTGCTAACCGTTTCACCTC
TAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCTAACGATGGTATCTCTCTGGCGCA
GACCACTGAAGGCGCACTGTCTGAGATTAACAACAACCTTACAACGTGTGCGTGAGTTGAC
TGTACAGGCGACCAACCGGTACTAACTCTGATTCTGACCTGGCTTCTATTTCAGGACGAAAT
CAAATCCCGTTTGTCTGAAATTGACCGCGTATCCGGGCAGACCCAGTTCAACGGCGTGAA
CGTATTGTCTAAAGATGGCTCCCTGAAAATTCAGGTTGGCGCAAATGATGGTCAGACTAT
CTCTATCGACCTGAAGAAAATTGACTCTGATACTCTGGGTTTGAATGGTTTCAACGTTAA
TGGTTCTGGTACCATTGCAAACAAAGCGGCCACAATCAGTGACTTGACTGCTCAGAAAGC
CGTTGACAACGGTAATGGTACTTATAAAGTTACAACCTAGCAACGCTGCACCTTACTGCATC
TCAGGCATTAAGTAAGCTGAGTGATGGCGTACTGTAGATATTGCAACCTATGCTGGTGG
TACAAGTTCAACAGTTAGTTATAAATACGACGCAGATGCAGGTAACCTCAGTTATAACAA
TACTGCAAACAAAACAAGTGCTGCGGCTGGAACCTCTGGCAGATACTCTTCTCCCGGCAGC
TGGCCAGACTAAAACCGGTACTTACAAGGCTGCTACTGGTGATGTTAACTTTAATGTTGA
CGCAACTGGTAATCTGACAATTGGCGGACAGCAAGCCTACCTGACTACTGATGGTAACCT
TACAACAAACAACCTCCGGTGGTGCGGCTACTGCAACTCTTAAAGAGCTGTTTACTCTTGC
TGCGGATGGTAAATCTCTGGGGAACGGCGGTACTGCTACCGTTACTCTGGATAATACTAC
GTATAATTTCAAAGCTGCTGCGAACGTTACTGATGGTGCTGGTGTCATCGCTGCTGCTGG
TGTAACCTTATACAGCCACTGTTTCTAAAGATGTCATTCTGGCACAACGCAATCTGCAAG
TCAGGCAGCAGCAACCGCTACCGACGGTGATACTGTGCAACGATCAACTATAAATCTGG
TGTCATGATCGGTTCCGCTACCTTTACCAATGGTAAAGGTACTGCCGATGGTATGACTTC
TGGTACAACCTCCAGTCGTAGCTACAGGTGCTAAAGCTGTATATGTTGATGGCAACAATGA
ACTGACTTCCACTGCATCTTACGATACGACTTACTCTGTCAACGCAGATACAGGCGCAGT
AAAAGTGGTATCAGGTACTGGTACTGGTAAATTTGAAGCTGTTGCTGGTGCGGATGCTTA
TGTAAGCAAAGATGGCAAATTAACGACAGAAACCACAGTGACGGCACTGCAACCAAAGA
TCCTTTGGCTGCCCTGGATGCTGCTATCAGCTCCATCGACAAATTCGGTTCCCTCCCTGGG
TGCTATCCAGAACCGTCTGGATTCCGCGAGTCACCAACCTGAACAACCACTACTAACCT
GTCTGAAGCGCAGTCCCGTATTTCAGGACGCCGACTATGCGACCGAAGTGCCAATATGTC
GAAAGCGCAGATCATCCAGCAGGCCGGTAACTCTGTGTTGGCAAAAGCTAACCGGTACC
GCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 54

77/96

ATGGCACAAGTCATTAATACCAACAGCC
TCTCGCTGATCACTCAAAATAATATCAACAAGAACCAGTCTGCGCTGTCGAGTTCTATCG
AGCGTCTGTCTTCTGGCTTGCGTATTAACAGCGCGAAGGATGACGCCGCAGGTCAGGCGA
TTGCTAACCGTTTTACTTCTAACATTAAAGGCCTGACTCAGGCTGCACGTAACGCCAACG
ACGGTATTTCTGTTGCACAGACCACTGAAGGCGCGCTGTCCGAAATCAACAACAATTAC
AGCGTATTCGTGAAGTACGCGTTTACGGCTTCTACCGGGACTAACTCTGATTTCGGATCTGG
ACTCCATTTCAGGACGAAATCAAATCCCGTCTCGACGAAATTGACCGCGTTTCCGGTCAAG
CCCAGTTCAACGGCGTGAACGTGCTGGCGAAAGACGGTTCGATGAAGATTTCAGGTTGGCG
CGAATGACGGGCAGACCATCTCTATCGATTTGCAGAAAATTGATTCTTCAACGCTGGGAT
TGAAAGGTTTCTCGGTATCAGGGAACGCATTAAAAGTTAGCGATGCGATAACTACAGTTC
CTGGTGCTAATGCTGGCGATGCCCCGTTACGGTTAAATTTGGTGCGAACGATACCGCTG
CTGCCGCAATGGCTAAAACATTGGGAATAAGTGATACATCAGGCTTGTCCCTACATAACG
TACAAAGCGCGGATGGTAAAGCGACAGGAACCTATGTTGTTCAATCTGGTAATGACTTCT
ATTCGGCTTCCGTTAATGCTGGTGGCGTTGTTACGCTTAATACCACCAATGTTACTTTCA
CTGATCCTGCGAACGGTGTACACAGCAACACAGACAGGTCAGCCTATCAAGGTCACGA
CGAATAGTGCTGGCGCGGCTGTTGGCTATGTTACTATTCAAGGCAAAGATTACCTTGCTG
GTGCAGACGGTAAGGATGCAATTGAAAACGGTGGTGACGCTGCAACAAATGAAGACACAA
AAATCCAACCTTACCGATGAAGTTCGATGTTGATGGTTCTGTAAAAACAGCGGCAACAGCAA
CATTTTCTGGTACTGCAACCAACGATCCGCTGGCACTTTTAGACAAAGCTATCTCGCAAG
TTGATACTTTCCGCTCCTCCCTCGGTGCCGTACAAAACCGTCTGGATTCTGCGGTCACCA
ACCTGAATAACACCACCACCAACCTGTCTGAAGCGCAGTCCCGTATTTCAGGACGCCGACT
ATGCGACCGAAGTGTCACCATGTCGAAAGCGCAGATCATCCAGCAGGCGGGTAACTCTG
TGCTGTCTAAAGCTAACCAGGTACCGCAGCAGGTTCTGTCTCTGCTGCAGGGTTAA

Figure 55

78/96

CTTCTCTTAGCTCTGCTATTGAGCGTCTGTCTTCTGGTCTGCGTATTAACAGCGCAAAAG
 ACGATGCAGCAGGTGAGGCGATTGCTAACCGTTTTACGGCAAATATTAAAGGTCTGACCC
 AGGCTTCCCGTAACGCGAATGATGGTATTTCTGTTGCGCAGACCACTGAAGGTGCGCTGA
 ATGAAATTAACAACAACCTGCAGCGTATTCGTGAACCTTCTGTTGAGGCAACTAACGGTA
 CTAACCTCTGACAGCGATCTTTCTTCTATCCAGGCTGAAATTACTCAACGTCTGGAAGAAA
 TTGACCGTGTATCTGAGCAAACTCAGTTTAAACGGCGTGAAAGTCCTTGCTGAAAATAATG
 AAATGAAAATTGAGTTGGTGCTAATGATGGTGAAACCATTGACCTGCCCCACGATTAG
 ATACAACACTCAGTTAGTAACGTGGAATCTTCATTCTCAGAATGACCCCTTCTCCAGCC
 CGCTGCAAATTCAGACGGTGTCTGATAATTCAGCGTGGAGTGCGGGCGGCATTGCTTATA
 ATCCTGCGCGCAGTCATTAATAATTTTCTGGCATGAACGATATCGCTGAACGAGTGCTC
 ATTCAAACATTTCATCGCGAAATCGTCCGTTAAAGCTCTCAATAAATCCGTTCTGCGTTGG
 CTTGCCCCGCTGGATTAAGCGCAACTCAACACCATGCTCAAAGGCCCATGATCCAGTGC
 ACGGCAAGTGAACTCCGGCCCTGGTCAGTTCTTATCGTCGCGGATAGCCTCGAAACAG
 TGCAATGCTGTCCAGAATACGCGTGACCTGAACGCGCTGAAATCCCAAAGGCAACAGTGAC
 CGTCAGGCATTCTTTGTGAAATCATCGACGAGGTAAGACACTTGATCCTGCGACCGGT
 GGAAAGTGCGTCCATGACGAAATCCATCGACCAGGTGAGTTGGGCGCCGCGGACGGAG
 CAGCGGCAGACGTTCTGTTGCCAGCCCTTTACGACGCTCTTCTGCGTTTTACGCCAGGCC
 ACTGAGGTGATAAAGCCGGTACACGCGCTTATGATTAACATGAAGCCCTTCACGGCGCAG
 CAACTGCCAAATACGACGGTAGCCAAAACGCGCTCGCTCCAGTGCCAGCTCAGTGATGCG
 CCCTGATAAATGCGCATCAGCAGCCGACGGTGAGCCTCATAGCGGCAGGTGACAGGGA
 TAAACCTGTAAGCCTGCAGGCACGACGTTGCGACAGACCGGTGCGATCACACATCAACAT
 CACGGCTTCCCGCTTCTGGTCTGTGCTCAGTACTTTGCCCCAGAGCCACCTGAAGCGCC
 TCTTTATCCAGCATGGCTTCGGCAAGCAGCTTCTTGAGTCTGGTGTCTCTCTCAAGC
 GACTTCAGGCGCTTAACCTTCAGGCACCTCCATACCGCCATACTTCTTACGCCAGGTGTAA
 AACGTGGCATCGGAAATGGCATGCTTGCGGCAGAGTTACAGGGCGGGTACCCAGCTTCG
 GCTTCGCGGAGAATACTGATGATCTGTTGCTCGGAAAAACGCTTCTTCATGGGGATGTCC
 TCATGTGGCTTATGAAGACATTACTAACATCGGGGTGTACTAATCAACGGGGAGCAGGTC
 ACCATCACTATCAATCTGGCAAAAATTGATGCGAAAACTCTCGGCCTGGACGGTTTTAAT
 ATCGATGGCGCGCAGAAAGCAACCGGCAGTGACCTGATTTCTAAATTTAAAGCGACAGGT
 ACTGATAATTATCAAATTAACGGTACTGATAACTATACTGTTAATGTAGATAGTGGAGTA
 GTACAGGATAAAGATGGCAACAAGTTTATGTGAGTGCTGCGGATGGTTCACCTACGACC
 AGCAGTGATACTCAATTCAAGATTGATGCAACTAAGCTTGCAAGTGGCTGCTAAAGATTTA
 GCTCAAGGTAATAAGATTGTCTACGAAGGTATCGAATTTACAAATACCGGCACTGGCGCT
 ATACCTGCCACAGGTAATGGTAAATTAACCGCCAATGTTGATGGTAAGGCTGTTGAATTC
 ACTATTTGCGGGAGTGCTGATACATCAGGTACTAGTGCAACCGTTGCCCTACGACAGCC
 CTATACAAAAATAGTGACGGGCAATTGACTGCAACAAAAGTTGAAAATAAAGCAGCGACA
 CTATCTGATCTTGATCTGAACGCTGCCAAGAAAACAGGAAGCACGTTAGTTGTTAACGGT
 GCAACTTACGATGTTAGTGCAGATGGTAAACGATAACGGAGACTGCTTCTGGTAACAAT
 AAAGTCATGTATCTGAGCAAATCAGAAGGTGGTAGCCCGATTCTGGTAAACGAAGATGCA
 GCAAAATCGTTGCAATCTACCACCAACCCGCTCGAAACTATCGACAAAGCATGGCTAAA
 GTTGACAATCTGCGTTCTGACCTCGGTGACGTACAAAACCGTTTCGACTCTGCCATCACC
 AACCTTGGCAACACCGTAAACAACCTGTCTTCTGCCCCGATGCCGTATCGAAGATGCTGAC
 TACGCGACCGAAGTGTCTAACATGTCTCGTGCGCAGATCTGCAACAAGCGGGTACCTCT
 GTTCTGGCACAGGCTAACC

Figure 56

79/96

AACAAAAACCAGTCTGCGCTGTGCGACTTCTATCGAGCGCCTCTCT
TCTGGTCTGCGCATTAAACAGCGCTAAAGATGACGCTGCGGGCCAGGCGATTGCTAACCGC
TTCACCTTCTAACATCAAAGGTCTGACTCAGGCCGCACGTAACGCCAACGACGGTATCTCT
CTGGCGCAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACCTGCAGCGTGTTCGT
GAACTGACCGTTCAGGCCACTACCGGTACTAACTCTGATTCTGACCTGTCTTCAATCCAG
GACGAAATCAAATCCCGTCTCGATGAAATTGACCGCGTATCCGGTCAGACTCAGTTCAAC
GGCGTGAAACGTACTGGCAAAAGATGGCTCGATGAAAATTCAGGTCGGTGCAAATGATGGT
CAGACAATCAGCATTGATTTGCAGAAGATTGATTCTTCTACTTTAGGGTTAAATGGTTTT
TCTGTTTCCAAAAATGCAGTATCTGTTGGTGATGCTATTACTCAATTGCCTGGCGAGACG
GCAGCCGATGCACCAGTAACCATCAAGTTTGATGATTGAGTAAAAACTGATTTAAACTG
ACCGATGCTTCAGGGTTAAGTCTGCATAACCTCAAAGATGAAAATGGTAATTTAACTAAC
CAGTATGTTGTACAGAATGGCGGAAAATCTTACGCTGCTACAGTCGCTGCCAATGGTAAT
GTTACGCTGAACAAAGCAAATGTAACCTACAGCGATGTCGCAAACGGTATTGATACCGCA
ACGCAGTCAGGCCAGTTAGTTTCAAGTTGGTGAGATTCTACCGGTACGCCAAAAGCATTC
GTGTCTGTCCAAGGTAAAAGCTTTGGCATTGATGACGCCGCCCTTGAAGAATAACACTGGT
GATGCTACCGCTACTCCACCGGGAACATCTGGGACAACAGTTGTGCGCAGCGTCAATTCAT
CTGAGTACGGGCAAAAACCTCTGTAGACGCTGATGTAACGGCTTCCACTGAATTCACAGGT
GCTTCAACCAACGATCCACTGACTCTGCTGGACAAAGCTATCGCATCTGTTGATAAATTC
CGTTCTTCTTTGGGGCGGTACAGAACCGTCTGAGCTCCGCTGTAACCAACCTGAACAAC
ACCACCACCAACCTGTCTGAAGCGCAGTCCCGTATTGAGGACGCCGACTATGCGACCGAA
GTGTCCAACATGTCGAAAGCGCAGATTATCCAGCAGGCAGGTAACCTCCGTGCTGTCCAAA

Figure 57

80/96

AACAAAAACCAGTCTGCGCTGTGCGACTTCTATCGAACGCCTCTCTTCTGG
CCTGCGTATTAAACAGTGCGAAAGATGACGCTGCCGGTCAGGCGATAGCTAACCGTTTCAC
CTCTAACATTAAAGGCCTGACTCAGGCTGCGCGTAACGCCAACGACGGTATTTCTCTGGC
GCAGACCACAGAAGGTGCGTTGTCTGAAATCAACAACAACCTGCAACGTGTGCGTGAGTT
GACCGTTCAGGCGACGACCGGTACTAACTCTGATTCTGACCTGTCATCTATTACAGGACGA
AATCAAATCCCGTCTGGATGAGATTGACCGTGTTCGCGTCAGACCCAGTTCAACGGCGT
GAATGTACTGGCAAAAGACGGTTCGATGAAGATTACAGTTGGCGCGAATGATGGCCAGAC
TATTAGCATTGATTACAGAAAATTGACTCTTCTACATTAGGGTTGAATGGTTTCTCCGT
TTCTGCTCAATCACTTAACGTTGGTGATTCAATTACTCAAATTACAGGAGCCGCTGGGAC
AAAACCTGTTGGTGTGATTTACTGCTGTTGCGAAAGATCTGACTACTGCGACAGGTAA
AACTGTGCGATGTTTCCAGCCTGACGTTACACAACACCCTGGATGCGAAAGGGGCTGCCAC
CGCACAGTTCGTCGTTCAATCCGGTAGTGATTCTACTCCGCGTCCATTGACCATGCAAG
TGGTGAAGTGACGTTGAATAAAGCCGATGTGCAATACAAAGACACCGATAATGGACTAAC
GACTGCAGCTACTCAGAAAGATCAGCTGATTAAAGTTGCCGCTGACTCTGACGGCGCGGC
TGCGGGATATGTAACATTCCAGGGTAAAACTACGCTACAACGGCTCCAGCGGCGCTTAA
TGATGACACTACGGCAACAGCCACAGCGAACAAGTTGTTGTTGAATTATCTACAGCAAC
TCCGACTGCGCAGTTCTCAGGGGCTTCTTCTGCTGATCCACTGGCACTTTTAGACAAAGC
CATTGCACAGGTTGATACTTTCCGCTCCTCCCTCGGTGCCGTTCAAACCGTCTGGACTC
TGCGGTAACCAACCTGAACAACACCACCACCAACCTGTCTGAAGCGCAGTCCCGTATTCA
GGACGCCGACTATGCGACCGAAGTGTCTAACATGTGAAAGCGCAGATCATCCAGCAGGC
GGGTAACCTCTGTGCTGTCTAAA

Figure 58

81/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG
AACCAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC
GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTAATTCTAA CATTAAAGGC
CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTTCTG TTGCACAGAC CACTGAAGGC
GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTTCT
ACCGGGACTA ACTCTGATTC GGATCTGGAC TCCATTTCAGG ACGAAATCAA ATCCCGTCTC
GACGAAATTG ACCGCGTATC CGGTCAGACC CAGTTCAACG GCGTGAACGT ACTGGCAAAA
GACGGTTCGA TGAAAATTCA GGTGCGTGC AATGACGGTG AAATATCAC TATCGACCTG
AAGAAAATCG ATTCTGATC TCTGGGTCTG AATGGTTTTA ACGTAAATGG TAAAGGTACT
ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAATTCTG CTGGCGCGAA GTTAAACAC
CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAATTACCG ATGCTGCATT
CGATAAATTA GGAATGCG ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAACGC
TAAATCTGGT GATTTTACTA CCACCAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC
GCAGGCTACT GATTCAGCTA AAAACGTGA TCGGTTAGCT GCCACCTTC ATGCTGATGT
GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA
TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT
GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAAGCCGC
GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAA
AGCAACTCCT GCGACAACCT CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTAATTATCA
GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAA GCGGCTGCCG CGACATCTTC
AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG
TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTG CCGACTATAC
AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC
GACTGATACC AATAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAATCCGC
GGGTAAATC ACTACTGAGA CTACCAGTGC TGGTTCTGCA ACGACCAACC CGCTTGCTGC
CCTGGACGAC GCTATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA
CCGTCTGGAT TCCGCAGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CTGAAGCGCA
GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTCC AACATGTCGA AAGCGCAGAT
TATCCAGCAG GCCGGTAACT CCGTGCTGGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT
GTCTCTGCTG CAGGGTTAA

Figure 59

82/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG
AACCAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC
GCGAAGGATG ACGCCGCAGG TCAGGCGATT GCTAACCGTT TTAATTCTAA CATTAAAGGC
CTGACTCAGG CGGCCCCGTA CGCCAACGAC GGTATTTCTG TTGCGCAGAC CACCGAAGGC
GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATTCGTG AACTGACGGT TCAGGCCACT
ACAGGGACTA ACTCCGATTC TGACCTGGAC TCCATCCAGG ACGAAATCAA ATCTCGTCTT
GATGAAATTG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT GCTGGCGAAA
GACGGTTCAA TGAAAATTCA GGTGGGTGCG AATGACGGCG AAACCATCAC GATCGACCTG
AAAAAATCG ATTCTGATAC TCTGGGTCTG AATGGCTTTA ACGTAAATGG TAAAGGTACT
ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAACCTCTG CTGGCGCGAA GTTAAACAC
CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACCTCCG ATGCTGCATT
CGATAAATTA GGAATGGCG ATAAAGTCAC AGTTGGCGGC GTAGATTATA CTTACAACGC
TAAATCTGGT GATTTTACTA CCACTAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC
GCAGGCTGCT GATTCAGCTT CAAAACGTGA TCGGTTAGCT GCCACCCTTC ATGCTGATGT
GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA
TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT
GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TCAAAGCAGC
GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACAGAATATA CCATCGCAAA
AGCAACTCCT GCGACAACCA CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTACTTATCA
GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC
AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG
TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATC ACTAACGTTG CCGACTATAC
AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC
GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAACCTCCG
GGGTAAAATC ACTACTGAGA CTACCAGTGC TGGTTCTGCA ACGACCAACC CGCTTGCTGC
CCTGGACGAC GCAATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA
CCGTCTGGAT TCCGCAGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CCGAAGCGCA
GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTCC AACATGTCCA AAGCGCAGAT
CATTACAGAG GCCGGTAACT CCGTGCTGGC AAAAGCTAAC CAGGTACCGC AGCAGGTTCT
GTCTCTGCTG CAGGGTTAA

Figure 60

83/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG
 AACCAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC
 GCGAAGGATG ACGCAGCGGG TCAGGCGATT GCTAACCGTT TTAATTCTAA CATTAAAGGC
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTTCTG TTGCGCAGAC CACCGAAGGC
 GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATTCGTG AACTGACGGT TCAGGCCACT
 ACAGGGACTA ACTCCGATTC TGACCTGGAC TCCATCCAGG ACGAAATCAA ATCTCGTCTT
 GATGAAATTG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT GCTGGCGAAA
 GACGGTTCAA TGAAAATTCA GGTGTTGTCG AATGACGGCG AAACCATCAC GATCGACCTG
 AAAAAAATCG ATTCCTGATAC TCTGGGTCTG AATGGCTTTA ACGTAAATGG TAAAGGTACT
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAACCTCTG CTGGCGCGAA GTTAAACAC
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACCTACC ATGCTGCATT
 CGATAAATTA GGAATGGCG ATAAAGTCAC AGTTGGCGGC GTAGATTATA CTTACAACGC
 TAAATCTGGT GATTTTACTA CCACTAAATC TACTGCTGGT ACGGTGTAA ACGCCGCGGC
 GCAGGCTGCT GATTCAGCTT CAAAACGTGA TCGTTAGCT GCCACCCTTC ATGCTGATGT
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA
 TTCAGCAGGT AATATACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TCAAAGCAGC
 GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACAGAATATA CCATCGCAAA
 AGCAACTCCT GCGACAACCA CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTACTTATCA
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC
 AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATC ACTAACGTTG CCGACTATAC
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGCACT GCTGTAAATG TGAACCTCCG
 GGGTAAAATC ACTACTGAGA CTACCAGTGC TGGTTCTGCA ACGACCAACC CGCTTGCTGC
 CCTGGACGAC GCAATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA
 CCGTCTGGAT TCCGCGGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CCGAAGCGCA
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTC AACATGTCGA AAGCGCAGAT
 CATCCAGCAG GCCGGTAACT CCGTGCTGGC AAAAGCTAAC CAGGTACCGC AGCAGGTTCT
 GTCTCTGCTG CAGGGTTAA

Figure 61

84/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG
AACCAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC
GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTA CTCTCTAA CATTAAAGGC
CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTTCTG TTGCACAGAC CACTGAAGGC
GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTTCT
ACCGGGACTA ACTCTGATTC GGATCTGGAC TCCATTACAG ACGAAATCAA ATCCCGTCTC
GACGAAATTG ACCGCGTATC CGGTCAGACC CAGTTCAACG GCGTGAACGT ACTGGCAAAA
GACGGTTCTG TGAAAATTCA GGTGTTGTCG AATGACGGTG AAATATCAC TATCGACCTG
AAGAAAATCG ATTCTGATAC TCTGGGTCTG AATGGTTTTA ACGTAAATGG TAAAGGTACT
ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAATTCTG CTGGCGCGAA GTTAAACAC
CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAATACCG ATGCTGCATT
CGATAAATTA GGGAATGGCG ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAACGC
TAAATCTGGT GATTTTACTA CCACCAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC
GCAGGCTACT GATTCAGCTA AAAAACGTGA TCGGTTAGCT GCCACCCTTC ATGCTGATGT
GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA
TTCAGCAGGT AATATCACC TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT
GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAAGCCGC
GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAAA
AGCAACTCCT GCGACAACCT CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTACTTATCA
GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC
AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG
TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTTG CCGACTATAC
AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC
GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAATCCGC
GGGTAAAATC ACTACTGAGA CTACCAAGTG TGGTTCTGCA ACGACCAACC CGCTTGCTGC
CCTGGACGAC GCTATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA
CCGTCTGGAT TCCGCAGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CTGAAGCGCA
GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTCC AACATGTCGA AAGCGCAGAT
TATCCAGCAG GCCGGTAACT CCGTGCTGGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT
GTCTCTGCTG CAGGGTTAA

Figure 62

85/96

ATGGCACAAGTCATTAATACCAACAGCCTCTCGCTGATCACTCAAATAATATCAACAAG
AACCAGTCTGCGCTGTGAGTTCTATCGAGCGTCTGTCTTCTGGCTTGCGTATTAACAGC
GCGAAGGATGACGCCGCGAGGTCAGGCGATTGCTAACCGTTTTACTTCTAACATTAAAGGC
CTGACTCAGGCGGCCCCGTAACGCCAACGACGGTATTTCTGTTGCGCAGACCACCGAAGGC
GCGCTGTCCGAAATCAACAACAACCTTACAGCGTATTCGTGAACTGACGGTTCAGGCCACT
ACAGGGACTAACTCCGATTCTGACCTGGACTCCATCCAGGACGAAATCAAATCTCGTCTT
GATGAAATTGACCGCGTATCCGGCCAGACCCAGTTCAACGGCGTGAACGTGCTGGCGAAA
GACGGTTCAATGAAAATTACAGTTGGTGCGAATGACGGCGAAACCATCACGATCGACCTG
AAAAAATCGATTCTGATACTCTGGGTCTGAATGGCTTTAACGTAAATGGTAAAGGTACT
ATTACCAACAAAGCTGCAACGGTAAGTGATTTAACCTCTGCTGGCGCGAAGTTAAACACC
ACGACAGGTCTTTATGATCTGAAAACCGAAAATACCTTGTTAACTACCGATGCTGCATT
GATAAATTAGGGAATGGCGATAAAGTCACAGTTGGCGGCGTAGATTATACTTACAACGCT
AAATCTGGTGATTTTACTACCACTAAATCTACTGCTGGTACGGGTGTAGACGCCGCGGCG
CAGGCTGCTGATTGAGCTTCAAAACGTGATGCGTTAGCTGCCACCCTTCATGCTGATGTG
GGTAAATCTGTTAATGGTTCTTACACCACAAAAGATGGTACTGTTTCTTTCGAAACGGAT
TCAGCAGGTAATATCACCATCGGTGGAAGCCAGGCATACGTAGACGATGCAGGCAACTTG
ACGACTAACACGCTGGTAGCGCAGCTAAAAGCTGATATGAAAGCGCTGCTCAAAGCAGCG
AGCGAAGGTAGTGACGGTGCTCTCTGACATTCAATGGCACAGAATATACCATCGCAAAA
GCAACTCCTGCGACAACCACTCCAGTAGCTCCGTTAATCCCTGGTGGGATTACTTATCAG
GCTACAGTGAGTAAAGATGTAGTATTGAGCGAAACCAAAGCGGCTGCCGCGACATCTTCA
ATTACCTTTAATTCCGGTGACTGAGCAAACTATTGGGTTTACCGCGGGTGAATCCAGT
GATGCTGCGAAGTCTTATGTGGATGATAAAGGTGGTATCACTAACGTTGCCGACTATACA
GTCTCTTACAGCGTTAACAAGGATAACGGCTCTGTGACTGTTGCCGGGTATGCTTCAGCG
ACTGATACCAATAAAGATTATGCTCCAGCAATTGGTACTGCTGTAAATGTGAACTCCGCG
GGTAAAATCACTACTGAGACTACCAGTGCTGGTTCTGCAACGACCAACCCGCTTGCTGCC
CTGGACGACGCAATCAGCTCCATCGACAAATTCCGTTCTTCCCTGGGTGCTATCCAGAAC
CGTCTGGATTCCGCGAGTCACCAACCTGAACAACACCACTACCAACCTGTCCGAAGCGCAG
TCCCGTATTACAGGACGCCGACTATGCGACCGAAGTGTCACACATGTCGAAAGCGCAGATC
ATTACAGCAGGCCGGTAACTCCGTGCTGGCAAAAGCTAACAGGTACCGCAGCAGGTTCTG
TCTCTGCTGCAGGGTTAA

Figure 63

86/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG
 AACCAGTCTG CGCTGTGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC
 GCGAAGGATG ACGCCGAGG TCAGGCGATT GCTAACCGTT TTAATTCTAA CATTAAAGGC
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTTCTG TTGCGCAGAC CACCGAAGGC
 GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATTCGTG AACTGACGGT TCAGGCCACT
 ACAGGGACTA ACTCCGATTC TGACCTGGAC TCCATCCAGG ACGAAATCAA ATCTCGTCTT
 GATGAAATTG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT GCTGGCGAAA
 GACGGTTCAA TGAAAATTCA GGTGGTGCG AATGACGGCG AAACCATCAC GATCGACCTG
 AAAAAAATCG ATTCTGATAC TCTGGGTCTG AATGGCTTTA ACGTAAATGG TAAAGGTACT
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAATTCTG CTGGCGCGAA GTTAAACAC
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAATTACCG ATGCTGCATT
 CGATAAATTA GGAATGGCG ATAAAGTCAC AGTTGGCGGC GTAGATTATA CTTACAACGC
 TAAATCTGGT GATTTTACTA CCACTAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC
 GCAGGCTGCT GATTCAGCTT CAAAACGTGA TGCCTTAGCT GCCACCCTTC ATGCTGATGT
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA
 TTCAGCAGGT AATATACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TCAAAGCAGC
 GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACAGAATATA CCATCGCAAA
 AGCAACTCCT GCGACAACCA CTCCAGTAGC TCCGTTAATC CCTGGTGGA TTAATTATCA
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAA GCGGCTGCCG CGACATCTTC
 AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATC ACTAACGTTG CCGACTATAC
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGCACT GCTGTAAATG TGAATCCGC
 GGGTAAATC ACTACTGAGA CTACCAGTGC TGGTCTGCA ACGACCAACC CGCTTGCTGC
 CCTGGACGAC GCAATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA
 CCGTCTGGAT TCCGCGGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CCGAAGCGCA
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTCC AACATGTCTGA AAGCGCAGAT
 CATCCAGCAG GCCGGTAACT CCGTGCTGGC AAAAGCTAAC CAGGTACCGC AGCAGGTTCT
 GTCTCTGCTG CAGGGTTAA

Figure 64

87/96

ATGGCACAAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG
 AACCAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC
 GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTAATTCTAA CATTAAAGGC
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTTCTG TTGCACAGAC CACTGAAGGC
 GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTTCT
 ACCGGGACTA ACTCTGATTC GGATCTGGAC TCCATTACAG ACGAAATCAA ATCCCGTCTC
 GACGAAATTG ACCGCGTATC CGGTACAGAC CAGTTCAACG GCGTGAACGT ACTGGCAAAA
 GACGGTTCGA TGAAAATTCA GGTGCGTGC AATGACGGTG AAATATCAC TATCGACCTG
 AAGAAAATCG ATTCTGATAC TCTGGGTCTG AATGGTTTTA ACGTAAATGG TAAAGGTACT
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAATTCTG CTGGCGCGAA GTTAAACACC
 ACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAATACCG ATGCTGCATT
 CGATAAATTA GGAATGGCG ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAACGC
 TAAATCTGGT GATTTTACTA CCACCAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC
 GCAGGCTACT GATTCAGCTA AAAAACGTGA TCGGTAGCT GCCACCCTTC ATGCTGATGT
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA
 TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAAGCCGC
 GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAA
 AGCAACTCCT GCGACAACCT CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTTCTTATCA
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAA GCGGCTGCCG CGACATCTTC
 AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTTG CCGACTATAC
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAATCCGC
 GGGTAAAATC ACTACTGAGA CTACCAAGTGC TGGTTCTGCA ACGACCAACC CGCTTGCTGC
 CCTGGACGAC GCTATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA
 CCGTCTGGAT TCCGCAGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CTGAAGCGCA
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTCC AACATGTCTG AAGCGCAGAT
 TATCCAGCAG GCCGGTAACT CCGTCTGGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT
 GTCTCTGCTG CAGGGTTAA

Figure 65

88/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG
AACCAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC
GCGAAGGATG ACGCCGCAGG TCAGGCGATT GCTAACCGTT TTAATTCTAA CATTAAAGGC
CTGACTCAGG CGGCCCGTAA CGCCAACGAC GGTATTTCTG TTGCGCAGAC CACCGAAGGC
GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATTCGTG AACTGACGGT TCAGGCCACT
ACAGGGACTA ACTCCGATTC TGACCTGGAC TCCATCCAGG ACGAAATCAA ATCTCGTCTT
GATGAAATTG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT GCTGGCGAAA
GACGGTTCAA TGAAAATTCA GGTGTTGTCG AATGACGGCG AAACCATCAC GATCGACCTG
AAAAAAATCG ATTCTGATAC TCTGGGTCTG AATGGCTTTA ACGTAAATGG TAAAGGTACT
ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAACCTCTG CTGGCGCGAA GTTAAACAC
CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACACCG ATGCTGCATT
CGATAAATTA GGGAAATGGCG ATAAAGTCAC AGTTGGCGGC GTAGATTATA CTTACAACGC
TAAATCTGGT GATTTTACTA CCACTAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC
GCAGGCTGCT GATTCAGCTT CAAAACGTGA TCGGTTAGCT GCCACCCTTC ATGCTGATGT
GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA
TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT
GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TCAAAGCAGC
GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACAGAATATA CCATCGCAAA
AGCAACTCCT GCGACAACCA CTCCAGTAGC TCCGTTAATC CCTGGTGGA TTACTTATCA
GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAA GCGGCTGCCG CGACATCTTC
AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG
TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATC ACTAACGTTG CCGACTATAC
AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC
GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAATCCGC
GGGTAAATC ACTACTGAGA CTACCAGTGC TGGTTCTGCA ACGACCAACC CGCTTGCTGC
CCTGGACGAC GCAATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA
CCGTCTGGAT TCCGAGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CCGAAGCGCA
GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTC AACATGTGCA AAGCGCAGAT
CATTCAGCAG GCCGGTAACT CCGTGCTGGC AAAAGCTAAC CAGGTACCGC AGCAGGTTCT
GTCTCTGCTG CAGGGTTAA

Figure 66

89/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG
 AACCAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC
 GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTAATTCTAA CATTAAAGGC
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTTCTG TTGCACAGAC CACCGAAGGC
 GCGCTGTCTG AAATCAACAA CAACTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTTCT
 ACCGGAACCTA ACTCTGATTC GGATCTGGAC TCCATTACAG ACGAAATCAA ATCCCGTCTT
 GATGAAATTG ACCGCGTATC CGGCCAGACC CAGTTCAACG GCGTGAACGT ACTGGCAAAA
 GACGGTTCCA TGAAAATTCA GGTGGTGCG AATGACGGTG AAATATCAC TATCGACCTG
 AAGAAAATCG ATTCTGATAC TCTGGGTCTG AATGGTTTTA ACGTAAATGG TAAAGGTACT
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAACCTCTG CTGGCGCGAA GTTAAACAC
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAACACCG ATGCTGCATT
 CGATAAATTA GGAATGGCG ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAACGC
 TAAATCTGGT GATTTTACTA CCACCAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC
 GCAGGCTACT GATTCAGCTA AAAACGTGA TCGGTTAGCT GCCACCCCTC ATGCTGATGT
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA
 TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAAGCCGC
 GAGCGAAGGT AGTGACGGTG CTTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAAA
 AGCAACTCCT GCGACAACCT CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTACTTATCA
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC
 AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTTG CCGACTATAC
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC
 GACTGATACC AATAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAATCCGC
 GGGTAAATC ACTACTGAGA CTACCAGTGC TGGTCTGCA ACGACCAACC CGCTTGCTGC
 CCTGGACGAC GCTATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA
 CCGTCTGGAT TCCGAGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CTGAAGCGCA
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTC AACATGTGCA AAGCGCAGAT
 TATCCAGCAG GCCGGTAACT CCGTGCTGGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT
 GTCTCTGCTG CAGGGTTAA

Figure 67

90/96

ATGGCACAAG TCATTAATAC CAACAGCCTC TCGCTGATCA CTCAAAATAA TATCAACAAG
 AACCAGTCTG CGCTGTCGAG TTCTATCGAG CGTCTGTCTT CTGGCTTGCG TATTAACAGC
 GCGAAGGATG ACGCCGCGGG TCAGGCGATT GCTAACCGTT TTAATTCTAA CATTAAAGGC
 CTGACTCAGG CTGCACGTAA CGCCAACGAC GGTATTTCTG TTGCACAGAC CACTGAAGGC
 GCGCTGTCCG AAATCAACAA CAACTTACAG CGTATCCGTG AGCTGACGGT TCAGGCTTCT
 ACCGGGACTA ACTCTGATTC GGATCTGGAC TCCATTACAG ACGAAATCAA ATCCCGTCTC
 GACGAAATTG ACCGCGTATC CGGTCAAGAC CAGTTCAACG GCGTGAACGT ACTGGCAAAA
 GACGGTTCGA TGAAAATTCA GGTGGGTGCG AATGACGGTG AAATATCAC TATCGACCTG
 AAGAAAATCG ATTCTGATAC TCTGGGTCTG AATGGTTTAA ACGTAAATGG TAAAGGTACT
 ATTACCAACA AAGCTGCAAC GGTAAGTGAT TTAATTCTG CTGGCGCGAA GTTAAACAC
 CACGACAGGT CTTTATGATC TGAAAACCGA AAATACCTTG TTAATACCG ATGCTGCATT
 CGATAAATTA GGGAAATGGCG ATAAAGTCAC CGTTGGCGGC GTAGATTATA CTTACAACGC
 TAAATCTGGT GATTTTACTA CCACCAAATC TACTGCTGGT ACGGGTGTAG ACGCCGCGGC
 GCAGGCTACT GATTCAGCTA AAAAACGTGA TCGGTTAGCT GCCACCCTTC ATGCTGATGT
 GGGTAAATCT GTTAATGGTT CTTACACCAC AAAAGATGGT ACTGTTTCTT TCGAAACGGA
 TTCAGCAGGT AATATCACCA TCGGTGGAAG CCAGGCATAC GTAGACGATG CAGGCAACTT
 GACGACTAAC AACGCTGGTA GCGCAGCTAA AGCTGATATG AAAGCGCTGC TTAAAGCCGC
 GAGCGAAGGT AGTGACGGTG CCTCTCTGAC ATTCAATGGC ACTGAATATA CTATCGCAAA
 AGCAACTCCT GCGACAACCT CTCCAGTAGC TCCGTTAATC CCTGGTGGGA TTTCTTATCA
 GGCTACAGTG AGTAAAGATG TAGTATTGAG CGAAACCAAA GCGGCTGCCG CGACATCTTC
 AATTACCTTT AATTCCGGTG TACTGAGCAA AACTATTGGG TTTACCGCGG GTGAATCCAG
 TGATGCTGCG AAGTCTTATG TGGATGATAA AGGTGGTATT ACTAACGTTG CCGACTATAC
 AGTCTCTTAC AGCGTTAACA AGGATAACGG CTCTGTGACT GTTGCCGGGT ATGCTTCAGC
 GACTGATACC AATAAAGATT ATGCTCCAGC AATTGGTACT GCTGTAAATG TGAATCCGC
 GGGTAAATC ACTACTGAGA CTACCAGTGC TGGTTCTGCA ACGACCAACC CGTTGCTGC
 CCTGGACGAC GCTATCAGCT CCATCGACAA ATTCCGTTCT TCCCTGGGTG CTATCCAGAA
 CCGTCTGGAT TCCGCGTCA CCAACCTGAA CAACACCACT ACCAACCTGT CTGAAGCGCA
 GTCCCGTATT CAGGACGCCG ACTATGCGAC CGAAGTGTC AACATGTGCA AAGCGCAGAT
 TATCCAGCAG GCCGGTAACT CCGTGCTGGC AAAAGCCAAC CAGGTACCGC AGCAGGTTCT
 GTCTCTGCTG CAGGGTTAA

Figure 68

91/96

ATGCGACGTATAGAACGAATACCGGGGTTATCGGCGTAAGCGGGGCAAA
GTTTACGATTTATTTTTTGGCTTAATGACACGAACAGCAACGAGGAAGGG
GAGTATTTTCGACCGCTAGAAAAAAATTCTAAAGGTTGTGAGTGACCAGAC
GATAACAGGGTTGACGGCGACGAAGCCGAAGGGTGGAAGCCCAATACTT
AAACCGTAGACTTGAAAACAGGAAAATGAATCATGGCACAAGTCATTAAT
ACCAACAGCCTCTCGCTGATCACTCAAAATAATATCAACAAGAACCAGTC
TGCGCTGTCGACTTCTATCGAGCGCCTCTCTTCTGGTCTGCGCATTAAACAG
CGCTAAAGATGACGCTGCGGGCCAAGCGATTGCTAACCGCTTCACTTCTA
ACATCAAAGGTCTGACTCAGGCCGCACGTAACGCCAACGACGGTATTTCT
CTGGCGCAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACCTTGCA
GCGTGTTTCGTGAACTGACCGTTCAGGCCACTACCGGTACTAACTCTGATTC
TGACCTGTCTTCAATACAGGACGAAATCAAATCCCGTCTCGATGAAATTG
ACCGCGTATCCGGTCAGACTCAGTTCAACGGCGTTAATGTTCTTTCCAAAG
ATGGTTCAATGAAAATTCAGGTTGGTGCGAATGATGGTCAAACATCTCC
ATCGATCTGAAGAAAATTGATTCTTCAACTTTGGGGCTGAATGGCTTCTCA
GTTTCTAAAACTCTCTTAATGTCAGCAATGCTATCACATCTATCCCGCAA
GCCGCTAGCAATGAACCTGTTGATGTTAACTTCGGTGATACTGATGAGTCT
GCAGCAATCGCAGCCAAATTGGGGGTTTCCGATACGTCAAGCCTGTCGCT
GCACAACATCCTTGATAAAGATGGTAAGGCAACAGCTGATTATGTTGTTT
AGTCAGGTAAAGACTTCTATGCTGCTTCTGTTAATGCCGCTTCAGGTAAAG
TAACCTTAAACACCATTGATGTTACTTATGATGATTATGCGAACGGTGTTG
ACGATGCCAAGCAAACAGGTCAGCTGATCAAAGTTTCAGCAGATAAAGAC
GGCGCAGCTCAAGGTTTTGTCACACTTCAAGGCAAAAACCTATTCTGCTGGT
GATGCGGCAGACATTCTTAAGAATGGAGCAACAGCTCTTAAGTTAACTGA
TCTGAATTTAAGTGATGTTACTGATACTAATGGTAAGGTAACCACAACCTGC
GACTGAGCAATTTGAAGGTGCTTCAACTGAGGATCCGCTGGCGCTTCTGG
ATAAAGCTATTGCATCAGTCGACAAATTCCGGTCTTCTCTAGGTGCCGTGC
AGAACCGTCTCGATTCCGCTATCACCAACCTGAACAACACCACCACCAAC
CTGTCTGAAGCGCAGTCCCGTATTCAGGACGCCGACTATGCGACCGAAGT
GTCCAACATGTCGAAAGCGCAGATCATCCAGCAGGCAGGTAACCTCCGTGC
TGTCTAAAGCGAACCAGGTACCGCAGCAAGTTCTGTCACTGTTACAAGGC
TAATGGCCTTAACCTGCCTGACCCCGCCACCGGCGGGGTTTTTCTGTCCG
CAATTTACCGATAACCCCCAAATAACCCCTCATTTACCCACTAATCGTCC
GATTAAAAACCTGCAGAAACGGATAATCATGCCGATAACTCATATAACG
CAGGGCTGTTTATCGTGAATTCACTCTATACCGCTGAAGGTGTAATGGATA
AACACTCGCTG

Figure 69

92/96

AACAGCCTCTCGCTGATCACTCAGAACAACATCAACAAAAACCACTCTTC
AATGTCTACTGCCATTGAGCGTCTGTCTTCCGGTCTGCGTATCAACAGCGC
AAAAGATGACGCTGCTGGCCAGGCGATTGCCAACCGCTTCACCTCTAACA
TCAAAGGTCTGACTCAGGCAGCTCGTAACGCCAACGACGGTATCTCCGTT
GCACAGACCACTGAAGGCGCACTGTCTGAAATCAACAACAACCTGCAGCG
TATCCGTGAGCTGACTGTTCACTCTTCTACGGGTACTAACTCTGAATCCGA
TCTGAACTCAATCCAGGACGAAATTAATCCCGTCTGGACGAAATTGACC
GCGTATCCGGTCAGACCCAGTTCAACGGCGTGAACGTGCTGGCAAAGAC
GGCTCCATGAAAATTCAGGTTGGCGCGAACGATGGTGAAACCATCACCAT
CGACCTGAAAAAATTGACTCTTCTACTTTAAACCTGACTGGGTTTAA

Figure 70A

93/96

CTCAGTATGCTGTCACCGGCAGTACAGGTGCCGTAAC TTACGATCCAGAT
ACAGATCCTGCCGCGACTGGTGATATTGTTTCTGCTTATGTTGATGATGCA
GGTACATTGACAACTGATGCAAACAAA ACTGTAAAATATTATGCCACAC
TAATGGTAGCGTCACGAACGACAGTGGTTCAGCTATTTACGCAACTGAAG
CGGGCAAATTGACTACTGAAGCGTCTACAGCTGCTGAAACTACCGCTAAC
CCACTGAAAGCCCTGGACGATGCAATCAGCCAGATCGACAAATTCGGTTC
TTCTCTGGGTGCTGTACAGAACCGTCTGGATTCTGCGGTAACCAACCTGAA
CAACACCACCACCAACCTGTCTGAAGCGCAGTCCCGTATTCAGGACGCCG
ACTATGCGACCGAAGTGTCAAATATGTCTAAAGCGCAGATCATCCAGCAG
GCCGGTAACTCCGTGTTGGCTAAAGCTAACCAGGTT CCTCAGCAGGTT

Figure 70B

94/96

AGCCTGTCGCTGTTGACCCAGAATAACCTGAACAAATCTCAGTCTTCTCTG
AGCTCCGCCATTGAGCGTCTCTCTTCTGGCCTGCGTATTAACAGTGCTAAA
GATGACGCAGCAGGTCAGGCGATTGCTAACCGTTTTACAGCAAATATTAA
AGGTCTGACTCAGGCTTCCCGTAACGCGAATGATGGTATTTCTGTTGCGCA
GACCACTGAAGGCGCGCTGAATGAAATTAACAACAACCTGCAGCGTGTAC
GTGAACTGACTGTTTCAAGGCAACTAACGGTACTAACTCTGACAGCGATCTTT
CTTCTATCCAGGCTGAAATTACTCAACGTCTGGAAGAAATTGACCGTGTAT
CTGAG CAAACTCAGTTTAACGGCGTGAAAGTCCTTGCTGAAAAT

Figure 71

95/96

GCACGTTAGTTGTTAACGGTGCAACTTACGATGTTAGTGCAGATGGTAAA
ACGATAACGGAGACTGCTTCTGGTAACAATAAAGTCATGTATCTGAGCAA
ATCAGAAGGTGGTAGCCCGATTCTGGTAAACGAAGATGCAGCAAAATCGT
TGCAATCTACCACCAACCCGCTCGAAACTATCGACAAAGCATTGGCTAAA
GTTGACAATCTGCGTTCTGACCTCGGTGCAGTACAAAACCGTTTCGACTCT
GCTATCACCAACCTTGGCAACACCGTAAACAACCTGTCTTCTGCCCGTAGC
CGTATCGAAGATGCTGACTACGCGACCGAAGTGTCTAACATGTCTCGTGC
GCAGATCCTGCAACAAGCGGGTACCTCTGTTCTGGCGCAGGCTAACCAGA
CCACGCAGAACGTAC

Figure 72

FIGURE 73A

Sequence of the polylinker region of plasmid pTrc99A:

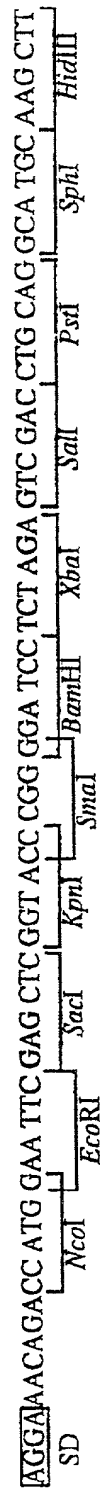
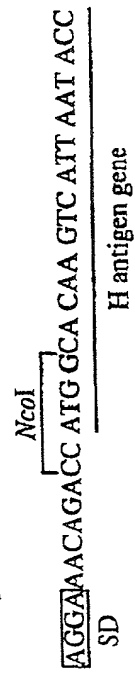


FIGURE 73B

Sequence in the junction region between vector and the 5' end of the H antigen gene:



DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I hereby declare:

That my residence, post office address and citizenship are as stated below next to my name.

That I verily believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural inventors are named below) of the subject matter which is claimed and for which a patent is sought on the invention entitled: **ANTIGENS AND THEIR DETECTION** the specification of which (check one)

☐ is attached hereto.

☒ was filed on 21/05/1999 as International Application Serial No. PCT/AU99/00385 and was amended on 01/05/2000 (if applicable).

That I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

That I acknowledge the duty to disclose information known to be material to patentability of this application in accordance with Title 37, Code of Federal Regulations §1.56(a).

That I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate on this invention having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s)

Priority Claimed

<u>PCT/AU99/00385</u> (Number)	<u>PCT</u> (Country)	<u>21/05/1999</u> (Day/Month/Year Filed)	<input checked="" type="checkbox"/> <input type="checkbox"/> Yes No
<u>PP3634</u> (Number)	<u>Australia</u> (Country)	<u>21/05/1998</u> (Day/Month/Year Filed)	<input checked="" type="checkbox"/> <input type="checkbox"/> Yes No
<u> </u> (Number)	<u> </u> (Country)	<u> </u> (Day/Month/Year Filed)	<input type="checkbox"/> <input type="checkbox"/> Yes No

That I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT international filing date of this application:

United States Application(s)

<u> </u> (Application Serial No.)	<u> </u> (Filing Date)	<u> </u> (Status)-(Patented, pending, abandoned)
<u> </u> (Application Serial No.)	<u> </u> (Filing Date)	<u> </u> (Status)-(Patented, pending, abandoned)
<u> </u> (Application Serial No.)	<u> </u> (Filing Date)	<u> </u> (Status)-(Patented, pending, abandoned)

That all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

I hereby appoint the following attorneys, with full power of substitution and revocation, to prosecute this application and to transact all business in the United States Patent and Trademark Office connected therewith and request that all correspondence and telephone calls in respect to this application be directed to: WELSH & KATZ, LTD., 120 South Riverside Plaza, 22nd Floor, Chicago, Illinois 60606-3913, Telephone No.: (312) 655-1500.

Attorney

Registration No.

15
A. Sidney Katz
Richard L. Wood
Jerold B. Schnayer
Eric C. Cohen
Joseph R. Marcus
Gerald S. Schur
Gerald T. Shekleton
James A. Scheer
Daniel R. Cherry
Kathleen A. Rheintgen
Edward P. Gamson
Thomas W. Tolpin
Shannon L. Nebolsky
Mitchell J. Weinstein
Charles R. Krikorian

24,003
22,839
28,903
27,429
25,060
22,053
27,466
29,434
29,054
34,044
29,381
26,700
41,217
37,963
40,687

Full name of sole or one
joint inventor:

Peter Richard Reeves

Inventor's signature:

P. Reeves

Date:

15/3/2001

Residence and Post Office Address:

20 Mansfield Street

Glebe NSW 2037

AUSTRALIA

Citizenship:

Great Britain

Full name of additional joint
inventor, if any:

Lei Wang

Inventor's signature:

Lei Wang

Date:

15/3/2001

Residence and Post Office Address:

8A Holt Street

North Ryde NSW 2113

AUSTRALIA

Citizenship:

Australia

Address for Correspondence:

WELSH & KATZ, LTD.
120 South Riverside Plaza
22nd Floor
Chicago, Illinois 60606-3913

PCT09

SV

RAW SEQUENCE LISTING

DATE: 10/23/2001

PATENT APPLICATION: US/09/701,132A

TIME: 10:29:44

Input Set : A:\P30384.app

Output Set: N:\CRF3\10232001\I701132A.raw

3 <110> APPLICANT: THE UNIVERSITY OF SYDNEY
 5 <120> TITLE OF INVENTION: ANTIGENS AND THEIR DETECTION
 7 <130> FILE REFERENCE: REEVES

C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/701,132A *over*

C--> 10 <141> CURRENT FILING DATE: 2001-06-15

ENTERED

12 <160> NUMBER OF SEQ ID NOS: 68

14 <170> SOFTWARE: PatentIn Ver. 2.0

16 <210> SEQ ID NO: 1

17 <211> LENGTH: 1773

18 <212> TYPE: DNA

19 <213> ORGANISM: Escherichia coli

21 <400> SEQUENCE: 1

```

22 atgcgacgta tagaacgaat accgggggta tcggcgtaag cggggcacaag ttacgattt 60
23 attttttggc ttaatgacac gaacagcaac gaggaagggg agtatttcga ccgctagaaa 120
24 aaaattctaa aggttgtag tgaccagacg ataacagggt tgacggcgac gaagccgaag 180
25 ggtggaagcc caatacttaa accgtagact tgaaaacagg aaaatgaatc atggcacaag 240
26 tcattaatac caacagcctc tcgctgatca ctcaaaaataa tatcaacaag aaccagtctg 300
27 cgtgtgtagc ttctatcgag cgctctctct ctggtctgag cattaacagc gctaaagatg 360
28 acgctgcggg ccaagcgatt gtaaacgct tcaattctaa catcaaaggt ctgactcagg 420
29 ccgcacgtaa cgccaacgac ggtatttctc tggcgagac cactgaaggc gactgtctg 480
30 aaatcaacaa caactgcag cgtgttctg aactgaccgt tcaggccact accggtacta 540
31 actctgattc tgacctgtct tcaatacagg acgaaatcaa atcccgctc gatgaaattg 600
32 accgcgtatc cggtcagact cagttcaacg gcgttaatgt tctttccaaa gatggttcaa 660
33 tgaaaattca ggttggtgag aatgatggct aaactatctc catcgatctg aagaaaattg 720
34 attcttcaac tttggggctg aatggcttct cagtttctaa aaactctctt aatgtcagca 780
35 atgctatcac atctatcccg caagccgcta gcaatgaacc tgttgatgtt aacttcgggtg 840
36 ataactgatg gtctgcagca atcgagcca aattgggggt ttccgatacg tcaagcctgt 900
37 cgctgcacaa catccttgat aaagatggta aggcaacagc tgattatgtt gttcagtcag 960
38 gtaaagactt ctatgctgct tctgttaatg ccgcttcagg taaagtaacc ttaaacacca 1020
39 ttgatgttac ttatgatgat tatgcgaacg gtgttgacga tgccaagcaa acaggtcagc 1080
40 tgatcaaagt ttcagcagat aaagacggcg cagctcaagg ttttgtcaca cttcaaggca 1140
41 aaaactattc tgctggtgat gcggcagaca ttcttaagaa tggagcaaca gctcttaagt 1200
42 taactgatct gaatttaagt gatgttactg atactaatgg taaggtaacc acaactgcga 1260
43 ctgagcaatt tgaagggtgt tcaactgagg atccgctggc gcttctggat aaagctattg 1320
44 catcagtcga caaattccgg tcttctctag gtgcgctgca gaaccgtctc gattccgcta 1380
45 tcaccaacct gaacaacacc accaccaacc tgtctgaagc gcagtcgccg attcaggacg 1440
46 ccgactatgc gaccgaagtg tccaacatgt cgaaagcgca gatcatccag caggcaggta 1500
47 actccgtgct gtctaaagcg aaccaggtag cgcagcaagt tctgtcactg ttacaaggct 1560
48 aatggcctta acctgctga ccccgccacc ggcgggggtt tttctgtccg caatttaccg 1620
49 ataaccacca aataaccctc catttcaccc actaatcgct cgattaaaaa cctgcagaa 1680
50 accgataatc atgccgataa ctcatataac gcagggtgtt ttatcgtgaa ttcactctat 1740
51 accgctgaag gtgtaatgga taaacactcg ctg 1773

```

53 <210> SEQ ID NO: 2

54 <211> LENGTH: 500

55 <212> TYPE: DNA

56 <213> ORGANISM: Escherichia coli

58 <400> SEQUENCE: 2

RAW SEQUENCE LISTING

DATE: 10/23/2001

PATENT APPLICATION: US/09/701,132A

TIME: 10:29:44

Input Set : A:\P30384.app

Output Set: N:\CRF3\10232001\I701132A.raw

```

59 aacagcctct cgctgatcac tcagaacaac atcaacaaaa accagtcttc aatgtctact 60
60 gccattgagc gtctgtcttc cggctctgct atcaacagcg caaaagatga cgctgctggc 120
61 caggcgattg ccaaccgctt cacctctaac atcaaaggct tgactcaggc agctcgtaac 180
62 gccaacgacg gtatctccgt tgcacagacc actgaaggcg cactgtctga aatcaacaac 240
63 aacctgcagc gtatccgtga gctgactgtt cagtcttcta cgggtactaa ctctgaatcc 300
64 gatctgaact caatccagga cgaaattaaa tcccgtctgg acgaaattga ccgcgtatcc 360
65 ggtcagaccc agttcaacgg cgtgaacgtg ctggcaaaaag acggctccat gaaaattcag 420
66 gttggcgcg aacgatggtg aaccatcacc atcgacctga aaaaaattga ctcttctact 480
67 ttaaactga ctgggtttaa 500

```

69 <210> SEQ ID NO: 3

70 <211> LENGTH: 500

71 <212> TYPE: DNA

72 <213> ORGANISM: Escherichia coli

74 <400> SEQUENCE: 3

```

75 ctcatgtatgc tgtcaccggc agtacagggt ccgtaactta cgatccagat acagatcctg 60
76 ccgcgactgg tgatattggt tctgcttatg ttgatgatgc aggtacattg acaactgatg 120
77 caaacaacac tgtaaaatat tatgccaca ctaatggtag cgtcacgaac gacagtgggt 180
78 cagctattta cgcaactgaa gcgggcaaat tgactactga agcgtctaca gctgctgaaa 240
79 ctaccgctaa cccactgaaa gccctggacg atgcaatcag ccagatcgac aaattccgtt 300
80 cttctctggg tgctgtacag aaccgtctgg attctgcggg aaccaacctg aacaacacca 360
81 ccaccaacct gtctgaagcg cagtcccgta ttcaggacgc cgactatgcg accgaagtgt 420
82 caaatatgtc taaagcgag atcatccagc aggcgggtaa ctccgtgttg gctaaagcta 480
83 accaggttcc tcagcaggtt 500

```

85 <210> SEQ ID NO: 4

86 <211> LENGTH: 399

87 <212> TYPE: DNA

88 <213> ORGANISM: Escherichia coli

90 <400> SEQUENCE: 4

```

91 agcctgtcgc tgttgaccca gaataacctg aacaaatctc agtcttctct gagctccgcc 60
92 attgagcgtc tctctctctg cctgcgtatt aacagtgtca aagatgacgc agcaggctcag 120
93 gcgattgcta accgttttac agcaaataat aaaggctctga ctcaggcttc ccgtaacgcg 180
94 aatgatggta tttctgttgc gcagaccact gaaggcgcgc tgaatgaaat taacaacaac 240
95 ctgcagcgtg tacgtgaact gactgttcag gcaactaacg gtactaactc tgacagcgat 300
96 ctttcttcta tccaggctga aattactcaa cgtctggaag aaattgaccg tgtatctgag 360
97 caaactcagt ttaacggcgt gaaagtcctt gctgaaaaat 399

```

99 <210> SEQ ID NO: 5

100 <211> LENGTH: 417

101 <212> TYPE: DNA

102 <213> ORGANISM: Escherichia coli

104 <400> SEQUENCE: 5

```

105 gcacgttagt tgttaacggt gcaacttacg atgttagtgc agatggtaaa acgataacgg 60
106 agactgcttc tggtacaacat aaagtcattg atctgagcaa atcagaaggt ggtagcccg 120
107 ttctggtaaa cgaagatgca gcaaaatcgt tgcaatctac caccaaccgg ctcgaaacta 180
108 tcgacaaaagc attggctaaa gttgacaatc tgcgttctga cctcgggtgca gtacaaaacc 240
109 gtttcgactc tgctatcacc aaccttggca acaccgtaaa caacctgtct tctgcccgta 300
110 gccgtatcga agatgctgac tacgcgaccg aagtgtctaa catgtctcgt gcgcagatcc 360
111 tgcaacaagc gggtaacctc gttctggcgc aggtctaacca gaccacgcag aacgtac 417

```

113 <210> SEQ ID NO: 6

114 <211> LENGTH: 950

RAW SEQUENCE LISTING

DATE: 10/23/2001

PATENT APPLICATION: US/09/701,132A

TIME: 10:29:44

Input Set : A:\P30384.app

Output Set: N:\CRF3\10232001\I701132A.raw

115 <212> TYPE: DNA

116 <213> ORGANISM: Escherichia coli

118 <400> SEQUENCE: 6

```

119 aacaaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgtatt 60
120 aacagcgcta aagatgacgc cgcgggccag gcgattgcta accgctttac ttctaacatc 120
121 aaaggtctga ctcaggccgc acgtaacgcc aacgacggta tttctctggc gcagacggct 180
122 gaaggcgcgc tgtcagagat taacaacaac ttgcagcgta ttcgtgaact gaccgttcag 240
123 gcctctaccg gcacgaactc tgattccgac ctgtcttcta ttcaggacga aatcaaatac 300
124 cgtcttgatg aaattgaccg tgtatctggt cagaccaggt tcaacgggtg gaacgtgctg 360
125 tcgaaaaacg attcgatgaa gattcagatt ggtgccaatg ataaccagac gatcagcatt 420
126 ggcttgcaac aaatcgacag taccactttg aatctgaaag gatttaccgt gtccggcatg 480
127 gcggatttca gcgcggcgaa actgacggct gctgatggta cagcaattgc tgctgcggat 540
128 gtcaaggatg ctgggggtaa acaagtcaat ttactgtctt acactgacac cgcgtctaac 600
129 agtactaaat atgcggctgt tgattctgca accggtaaat acatggaagc cactgtagtc 660
130 attaccggtg cggcggcggt ggtaactggt ggtgcagcgg aagtggcggg agccgctaca 720
131 gccgatccgt taaaagcact ggatgccgca atcgctaaag tcgacaaatt ccgtcctcc 780
132 ctcggtgccg ttcaaaaccg tctggattct gcggtcacca acctgaacaa caccaccacc 840
133 aacctgtctg aagcgcagtc ccgtattcag gaagccgact atgcgaccga agtgtccaac 900
134 atgtcgaaag cgcagattat ccagcaggcg ggcaactccg tgctgtctaa 950

```

136 <210> SEQ ID NO: 7

137 <211> LENGTH: 1212

138 <212> TYPE: DNA

139 <213> ORGANISM: Escherichia coli

141 <400> SEQUENCE: 7

```

142 aacaaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgtatt 60
143 aacagcgcta aagatgacgc cgcgggccag gcgattgcta accgcttcac ttctaacatc 120
144 aaaggtctga ctcaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180
145 gaaggcgcgc tgtctgaaat caacaacaac ttgcagcggt tgcgtgagtt gaccgttcag 240
146 gcgacgaccg ggactaactc tgattctgac ctgtcttcta ttcaggacga aatcaaatac 300
147 cgtctggatg aaattgatcg cgtttccggt cagaccaggt tcaacggcgt gaatgtgctg 360
148 gcgaaagatg gttcgatgaa gattcaggtt ggcgcgaatg atgggcagac tattagcatt 420
149 gatttgcaga agattgactc ttctacatta ggactgaacg gtttctccgt ttcgggtcag 480
150 tcaacttaacg ttagtgattc cattactcaa attaccggtg ccgccgggac aaaacctggt 540
151 ggtgttgatt tcaactgctg tgcgaaagat ctgactactg cgacaggtaa aacagtcgat 600
152 gtttctagcc tgacgttaca caacactctg gatgcgaaag gggctgctac atcacagttc 660
153 gtcgttcaat ccggcaatga tttctactcc gcgtcgatta atcatacaga cggcaaagtc 720
154 acgttgaaata aagccgatgt cgaatacaca gacaccgata atggactaac gactgcggct 780
155 actcagaaag atcaactgat taaagttgcc gctgactctg acggctcggc tcggggatat 840
156 gtaacattcc aaggtaaaaa ctacgtaca acggtttcaa cggcacttga tgataatact 900
157 gcggcaaaag caacagataa taaagttggt gttgaattat caacagcaa accgactgca 960
158 cagttctcag gggcttcttc tgctgatoca ctggcacttt tagacaaagc tattgcacag 1020
159 gttgatactt tccgctcctc cctcggtgcg gtgcaaaacc gtctggattc cgcagtaaac 1080
160 aacctgaaca acaccaccac caacctgtct gaagcgcagt cccgtattca ggacccgac 1140
161 tatgtacag aagtgtccaa catgtcgaaa gcgcagatca tccagcaggc aggtaaactc 1200
162 gtgctgtcca aa 1212

```

164 <210> SEQ ID NO: 8

165 <211> LENGTH: 1647

166 <212> TYPE: DNA

167 <213> ORGANISM: Escherichia coli

RAW SEQUENCE LISTING

DATE: 10/23/2001

PATENT APPLICATION: US/09/701,132A

TIME: 10:29:44

Input Set : A:\P30384.app

Output Set: N:\CRF3\10232001\I701132A.raw

169 <400> SEQUENCE: 8

```

170 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
171 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
172 gcgaaggatg acgccgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaagggc 180
173 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
174 gcgctgtccg aaattaacaa caacttacag cgtattcgtg aactgacggt tcaggcttct 300
175 accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgctc 360
176 gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
177 gacggttcga tgaaaattca ggttggtgcg aatgacggcc agactatcac tattgatctg 480
178 aagaaaattg actctgatac gctggggctg aatgggttta atgtgaacgg caaaggggaa 540
179 acggctaata cggcagcaac cctgaaagat atgtctggat tcacagctgc ggcgccacca 600
180 gggggaactg ttggtgtaac tcaatatact gacaaatcgg ctgtagcaag tagcgtagat 660
181 attctaaatg ctgttctggt cgcagatgga aataaagtta caactagcgc cgatgttggt 720
182 tttggtacac cagccgtgc tgtaacctat acctacaata aagacactaa ttcatttcc 780
183 gccgcttctg atgatatctc cagcgctaac ctggctgctt tcctcaatcc tcaggccgga 840
184 gatacgacta aagctacagt tacaattggt ggcaaagatc aagatgtaaa catcgataaa 900
185 tccggtaatt taactgctgc tgatgatggc gcagtaactt atatggatgc taccgtaaac 960
186 ttaactaaaa ataatgctgg tgggtgataca caagctactt tggctaaact tgctactgct 1020
187 actggtgcta aagccgcgac catccaaact gataaaggaa cattcaccag tgacggtaca 1080
188 gcgtttgatg gtgcacaaat gtccattgat accaatacat ttgcaaatgc agtaaaaaat 1140
189 gacacttata ctgccactgt aggtgctaag acttatagcg taacaacagg ttctgctgct 1200
190 gcagacaccg cttatatgag caatggggtt ctcaagtata ctccgccaac ttactatgca 1260
191 caagctgatg gaagtatcac aactactgag gatgcggctg ccggtaaact ggtctacaaa 1320
192 ggttccgatg gtaagttaac aacggatacg actagcaaa gagaatcaac atcagatccg 1380
193 ctggcagctc ttgacgacgc tatcagccag atcgacaaat tccgctcctc cctgggtgcg 1440
194 gtgcaaaacc gtctggattc cgcagtgacc aacctgaaca acaccactac caacctgtct 1500
195 gaagcgcagt cccgtattca ggacgccgac tatgcgaccg aagtgtccaa catgtcgaaa 1560
196 gcgcagatta tccagcaggc cggtaactcc gtgctggcaa aagctaacca ggttccgcag 1620
197 caggttctgt ctctgctgca gggtaa 1647

```

199 <210> SEQ ID NO: 9

200 <211> LENGTH: 1758

201 <212> TYPE: DNA

202 <213> ORGANISM: Escherichia coli

204 <400> SEQUENCE: 9

```

205 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
206 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
207 gcgaaggatg acgccgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaagggc 180
208 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac caccgaaggc 240
209 gcgctgtctg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggcttct 300
210 accggaacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgctt 360
211 gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt actggcaaaa 420
212 gacggttcga tgaaaattca ggttggtgcg aatgacgggt aaactatcac tatcgacctg 480
213 aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaaatgg taaaggtaact 540
214 attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacacc 600
215 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
216 gataaattag ggaatggcga taaagtcacc gttggcggcg tagattatac ttacaacgct 720
217 aaatctggtg attttactac caccaaatct actgctggta cgggtgtaga cggcgcgcg 780
218 caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
219 ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900

```

RAW SEQUENCE LISTING

DATE: 10/23/2001

PATENT APPLICATION: US/09/701,132A

TIME: 10:29:44

Input Set : A:\P30384.app

Output Set: N:\CRF3\10232001\I701132A.raw

```

220 tcagcaggta atatcaccat cgggtgaagc caggcatacg tagacgatgc aggcaacttg 960
221 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgctgct taaagccgcg 1020
222 agcgaaggta gtgacggtgc ttctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
223 gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
224 gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
225 attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
226 gatgctgcga agtcttatgt ggatgataaa ggtggtatta ctaacgttgc cgactataca 1320
227 gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
228 actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440
229 ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
230 ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccttgggtgc tatccagaac 1560
231 cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgag 1620
232 tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgagatt 1680
233 atccagcagg ccggtaaact cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
234 tctctgctgc agggtaa

```

236 <210> SEQ ID NO: 10

237 <211> LENGTH: 1383

238 <212> TYPE: DNA

239 <213> ORGANISM: Escherichia coli

241 <400> SEQUENCE: 10

```

242 aacaaatctc agtcttctct tagctctgct attgagcgct tgtcttcttg tctgcgtatt 60
243 aacagcgcaa aagacgatgc agcaggtcag gcgattgcta accgttttac ggcaaatatt 120
244 aaaggctctga cccaggcttc ccgtaacgca aatgatggta tttctgttgc gcagaccact 180
245 gaagggtgcgc tgaatgaaat taacaacaac ctgcagcgta ttcgtgaact ttctgttcag 240
246 gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggctga aattactcaa 300
247 cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360
248 gctgaaaata atgaaatgaa aattcagggt ggtgctaatt atggtgaaac catcactatc 420
249 aatctggcaa aaattgatgc gaaaactctc ggccctggac gttttaatat cgatggcgcg 480
250 cagaaagcaa caggcagtga cctgatttct aaattttaaag cgacaggtac tgataattat 540
251 gatgttggcg gtaaaactta taccgtgaat gtggagagcg gcgcggttaa gaatgatgct 600
252 aataaagatg tttttgtaag cgcagctgat ggatcgctga cgaccagtag tgatactaaa 660
253 gtatccggtg aaagtattga tgcaacagaa ctagcgaaac ttgcaataaa attagctgac 720
254 aaaggctcca ttgaatacaa gggcattaca tttactaaca acactggcgc agagcttgat 780
255 gctaattgta aagggtgtttt gaccgcaaat attgatggtc aagatgttca atttactatt 840
256 gacagtaatg caccacggg tgccggcgca acaataacta cagacacagc tgtttacaaa 900
257 aacagtgcgg gccagttcac cactacaaaa gtggaaaata aagccgcaac actctctgat 960
258 ctggatctta atgcagccaa gaaaacaggt agcactttag ttgtaaatgg cgccacctac 1020
259 aatgtcagcg cagatggtaa aacggtaaact gatactactc ctggtgcccc taaagtgatg 1080
260 tatctgagca aatcagaagg tggtagcccg attctggtaa acgaagatgc agcaaaatcg 1140
261 ttgcaatcta ccaccaaccc gctcgaaact atcgacaagg cattggctaa agttgacaat 1200
262 ctgcgttctg acctcggtgc agtacaaaac cgtttcgact ctgccatcac caaccttggc 1260
263 aacaccgtaa acaacctgtc ttctgcccggt agccgtatcg aagatgctga ctacgcgacc 1320
264 gaagtgtcta acatgtctcg tgcgcagatc ctgcaacaag cgggtacctc tgttctggcg 1380
265 cag

```

267 <210> SEQ ID NO: 11

268 <211> LENGTH: 2013

269 <212> TYPE: DNA

270 <213> ORGANISM: Escherichia coli

272 <400> SEQUENCE: 11

DATE: 10/23/2001

TIME: 10:29:47

Output Set: N:\CRF3\10232001\I701132A.raw

L:9 M:270 C: Current Application Number differs, Replaced Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

[illegible]

PCT09

RAW SEQUENCE LISTING

PATENT APPLICATION US/09/701,132

DATE: 01/23/2001

TIME 11.14 28

Input Set A:\P30384.app

Output Set N:\CRF3\01232001\I701132.raw

ENTERED

5000

3 <110> APPLICANT THE UNIVERSITY OF SYDNEY
 5 <120> TITLE OF INVENTION: ANTICENS AND THEIR DETECTION
 7 <130> TITLE REFERENCE PFEVES

C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/701,132

C--> 10 <141> CURRENT FILING DATE: 2000-11-21

12 <160> NUMBER OF SEQ ID NOS 68

14 <170> SOFTWARE PatentIn Ver 2.0

16 <210> SEQ ID NO 1

17 <211> LENGTH 1773

18 <212> TYPE DNA

19 <213> ORGANISM Escherichia coli

21 <400> SEQUENCE 1

22 atgcacgca tagaacgaat accggggtta tcggcglaag cggggcaaa ttacgattt 60
 23 atttttggc ttaatgacac gaacagcaac gaggaaagg agtatttoga ccgctagaaa 120
 24 aaaaattctaa aggttgtgag tgaccagacg ataacagggt tgacggcgac gaagccgaaq 180
 25 ggtgcaagcc caatacttaa accgtggact tgaadacag aaaaatgaat atggcacaac 240
 26 tcaatcaaac caacagcccc tccctgatac ctcaaaaata taacaacaay aaccagttctg 300
 27 cgtgtgcac ttctatcgag cgcctctctt ctggtctgcg catcaaacg gctaaagatg 360
 28 acgctgggg ccaagcgatt gctaacogct caacttctaa calcaaaagg ctgactcagg 420
 29 ccgcacgtaa cgcacacgac ggtatttctc tggcgacag caacgaagge gcactgtctg 480
 30 aaaaacaaa caacttgcag cgtgttctgt aactgaacgt tcaagccact accgactcta 540
 31 actctgctc tgactctctt tcaatcacgg acgaatcga atccgtctc gatgaattg 600
 32 accgctatc cggctcagact cagttcaacg gcttaattgt tctttccaaa gatggttcaa 660
 33 tgaadacaa ggttgggtgc aatgatggtt aaactatctc cactgatctg aagaaaaattg 720
 34 attctcaac ttggggctcg aatggtttct cagtttctaa aaactctctt aatgtcagca 780
 35 atgtatcac atctatcccg caaacccgca gcaatcaacc tgttgatgtt aacttcggtg 840
 36 atactgacga gtttgacgca atcgagccca aatgggggtt ttccgatacg caagccctgt 900
 37 cgtgcacaa cactcctgac aaagctggta aggcacacac tgatttatgt tttcagttca 960
 38 gtaaaagact ccatgctgct tctgttaatg ccgcttcagg taagtaaac taacaacca 1020
 39 ttgatgttac ttatgatgat tatgcgaacg atgttganga tgcacagcaa acaggtcagg 1080
 40 tgatcaaat ttacagcgat aaagacggcg cagctcaagg ttgtgtaaca ctcaagga 1140
 41 aaaaactatc tctgtgtgat ggcgcagaca ttcttaagaa tggagcaaca gctcttaagt 1200
 42 taactgact gaatttaagt gatgttcttg atactaatgg taaggtaacc acaactgcga 1260
 43 ctgagcaatt tgaaggtgct tcaactgagg atccgctggc gcttctggat aaagctattg 1320
 44 cactagtcga caaatccgg tcttctctag gtgccttgca gaacccgtct gattccgcta 1380
 45 tcaaccaact gaacaaac accaccaacc tgtctgaagc gcagtcacct attcaggacg 1440
 46 ccgaactatc gaccgaagtg tcaaacatgt cgaagcgca gatcatccag caggcaggtt 1500
 47 actccgtgct gttcaaaagc aacraagttac cgaagcaagt tcgtgactg ttcaaggtct 1560
 48 aatggctta acctgcttga cccgcacac ggcggggtt ttctgttcg caatttaacg 1620
 49 ataaacccca aataaacctt catctcacc actaatgctc cgttcaaaa ccttcagaa 1680
 50 accgataatc atgcagata c caataaac gcagggctgt ttatcgtgaa ttcaactctat 1740
 51 accgtgaag gttgtaatgga taacacactg ctg 1773

53 <210> SEQ ID NO 2

54 <211> LENGTH 500

55 <212> TYPE DNA

56 <213> ORGANISM Escherichia coli

58 <400> SEQUENCE 2

RAW SEQUENCE LISTING
 PATENT APPLICATION US/09/701,132
 DATE 01/23/2001
 TIME 12 14 28

Input Set A:\P30384.app
 Output Set N:\CRF3\01232001\I701132.raw

```

59 aacagcctct cgtgcaaac tgggacaaac atcaacaaac aacagctc aatgtctact 60
60 gccattgagc gctgtctctt cgtcttggtt atcaacagcg caaaagatga cgtgtctggc 120
61 cagcgcattg ccaacgcctt cactctaac atcaaaagtc tgaatcagge aatcgtgac 180
62 gccaacgagc gtatctcgtt tgcacagac actgaagcgc cactgtctg aatcaaacac 240
63 aactctgagc gtatcctgta gctgacgtt cgtctctcta cgggtactaa cctgcaatcc 300
64 gatctgaact caatccagga cgaattaaa tcccgctggc atgaattga cgcgtatcc 360
65 ggtcagaccc agttcaacgc cgtgaacgtg ctggcacaag acgctcccat gaaattccag 420
66 gttagcgcga atgattggtg aacatccacc atcgactga aaaaattga cctctctact 480
67 ttatccctga ctgggtttaa
68
69 <210> SEQ ID NO 3
70 <211> LENGTH 500
71 <212> TYPE DNA
72 <213> ORGANISM Escherichia coli
73 <400> SEQUENCE 3
74
75 ctcaatgagc tgcacgcgc agtacagctg cgttaacta cgtatccagt acatctctg 60
76 cgcgcagctg tgaattgtt tctgttatg ttgatgatg aggtacattg acaactgat 120
77 caaacaaac tgaataat tctccacc ctaattggtg cgtcagcaac gacatgtgt 180
78 cagctatgta cgcacatgaa cgcgcacaa tgaactacta agcgtctaca cgtgtctgaa 240
79 ctacgcctaa cccactgaaa cgcctgagc atgaatcag ccagatcgac aattctcgt 300
80 ctctctctgg tgcgtacag aacgcctctg attctgcgtt aaccaactg aacaacacca 360
81 ccacaaact gctgagcgc cgtctcccta ttcagagcgc cgtatctgag ccggaagtgt 420
82 caaatatgta taagcgccag atcatccagc aggcgcgtaa ctccgtgttg gctaaagcta 480
83 aacagcttcc tgcagcgtt
84
85 <210> SEQ ID NO 4
86 <211> LENGTH 399
87 <212> TYPE DNA
88 <213> ORGANISM Escherichia coli
89 <400> SEQUENCE 4
90
91 agcctgtcgc tgttgcacca gaataacctg aacaaatctc agtctctctt gactctgcc 60
92 attgacgcgc tctctctctg cctgcgtatt aacagtgcta aagatgagc agcaggtcag 120
93 gcgattgcta accgcttctc atcaaatatt aaaaacttga ctacagcttc cgttaacgc 180
94 aatgattgta tctctcttgc gacagcactt caagggcgcg tgaatgaaat taacaacaa 240
95 ctgcagcgtg taactgaact gactgtcag gcaactaac gtaactactc tgcagcgtat 300
96 ctctctctga tccagcttga aattatcaa cgtctggaag aatttgacc tgtatctgag 360
97 caaatcactt taaagcgtt gaaagctctt gctgaaat
98
99 <210> SEQ ID NO 5
100 <211> LENGTH 417
101 <212> TYPE DNA
102 <213> ORGANISM Escherichia coli
103 <400> SEQUENCE 5
104
105 gacagctagt tgttaacgtt gcaacttacc atcttctgac agatggtaaa acgataacgc 60
106 agactgcttc tggtaacuat aaagtcatgt atctgagcga atcaagaggt gtagcccca 120
107 tctctgtata cgaagatgca gcaaatctgt tgaatctac caacacccg ctgaaacta 180
108 tgcacaaagc attgctaaa gttagacac tgcgtctga cctcggctga gtacaaac 240
109 gtttcgactc tctatctacc aactttgcca acaccgtaaa caacctgtct tctgccccta 300
110 gcgctatcga agatgtgac tgcagcaccg aagtgtctaa cagtctctgt ggcagatcc 360
111 tgcacaaagc ggttaactct gttctggcgc aggttaacca gacacgcag aactgac 417
112
113 <210> SEQ ID NO 6
114 <211> LENGTH 950

```

RAW SEQUENCE LISTING
 PATENT APPLICATION US/09/701,132 DATE 01/23/2001
 TIME 12:14:29

Input Set A:\P30384.app
 Output Set N:\CRF3\01232001\I701132.raw

```

115 <212> TYPE: DNA
116 <213> ORGANISM: Escherichia coli
118 <400> SEQUENCE: 6
119 aacaaaaacc agtctagcgt atcgacttct atcgagcgcc tctcttctgg tctgcgtatt 60
120 aacagcgcta aagatgacgc cggggacacg gggattgcta acgcctttac ttctaacatc 120
121 aaaggtctga ctccagcgcc acg'aacgcc aacgacggta ttctcttgge gcagacggct 180
122 gaagcgcgcc tctcagcaat taacaaacac ttacagcgta ttctggaact gacgcgttca 240
123 ggccttaccg gacagacgac caattccgac ctctcttcta ttacggacga aatcaaatcc 300
124 cctcttgatg aaatlgacgg tctatctggc cagaccagc tcacgggtg gaacgtgctg 360
125 tccaaaaacg attcgatgaa gat'cagat'c gctgcgaatg ataacccagc gacacagatt 420
126 ggtctgacac aaatcgacag taacacttgc aatctgaaag gatttaccgt gtcggcgatg 480
127 ggcgatttca ggcggcgcaa actgacggct gctgattgta cagcaattgc tgcctgcgat 540
128 gtcagggtg ctgggggtaa acaagtcac ttactgctt acatgacac cgcgtctaac 600
129 aatctcaaat atgcgggtct tga'cttgcga acgggtcaat acatgcaagc cactgtatgc 660
130 attacccgta cggcgcgccc ggttaacty r agtgcacgc aagtggcgcc agcgcctaca 720
131 ggcgattcgt taaacacat ggcacccgca atcgctaaag tcgacaaatt cgcctctccc 780
132 ctgcgtgccc tcaaaaacgc actggattct ggcgtcacca acctgaacaa caccaccacc 840
133 aacat'gtctg aagcgcajtc cctatttcaa gacgcgcact atgcgaccca agtctccaac 900
134 a'gtcgaagc ggcagattat ccagcagcg ggcacactcc tgcgtcttaa 960
136 <210> SEQ ID NO: 7
137 <211> LENGTH: 1212
138 <212> TYPE: DNA
139 <213> ORGANISM: Escherichia coli
141 <400> SEQUENCE: 7
142 aacaaaaacc agtctagcgt atcgacttct atcgagcgcc tctcttctgg tctgcgtatt 60
143 aacagcgcta aagatgacgc cggggacacg gggattgcta acgccttcaac ttctaacatc 120
144 aaaggtctga ctccagcgcc acat'aacgcc aacgacggta tctctctggc gcagacacac 180
145 gaagcgcgcc tctcagcaat taacaaacac ttacagcgta ttctggaact gacgcgttca 240
146 ggcagcagcc ggcactaacct cgaattctgc ctctcttcta ttacggacga aatcaaatcc 300
147 cgtctagatg aaatlgatcg ggtttccgg' cagacccagt tcaacggcgt gac'gtctg 360
148 ggcgaagatg gtlcgatgaa gat'cagatt' ggcgcgaatg atggcgcaac tat'aagcatl 420
149 gatttgcaga agattgacac tctacattta ggcctgaaag gtttctccgt ttccggctca 480
150 tcaacttaacg ttagtgaatc cat'actcaaa at'acgggtg cggcggggac aacacctggt 540
151 ggtgtraatc tcaactgctg' tgcgaagat ctgaactact cgcacggtaa aacagtcgat 600
152 gtttctagcc tgaacgtaca caaacac'ctg gatgcgaag ggcctactac atcacagtlc 660
153 gtcgttcaat cggcgaatga ttctacttcc agctcgatta atcatacaga cggcaaaagt 720
154 acgttgaata aagcagatg' cgaatcacca gacacgata atgactaac gactgcgct 780
155 ac'cagaaag atcactgat' taag'ctgc' cctgactctg acggtctggc tgcgggatat 840
156 gtaacattcc a'ggtaaaaa ctacgc'aca acggtttcaa cggcacttga' gataataact 900
157 ggcgcaaaag caacagataa taagltggt' gttgaabrat caacagcaaa acgcactgca 960
158 cagttctcag ggccttcttc tgcgtatccc ct'ggcacttt taacaaaagc tattgcacag 1020
159 gttgataact tccgtctctc cctcgctcgc gtcgaacacc gtcggatcc cgcagtaacc 1080
160 acctgaac'c acacacccc caacctgtct gaagcgcagt cccgtattca ggaacgcac 1140
161 tatgtacag aagtgctcaa catglogaaa ggcagatca tccagcagge aggttaactc 1200
162 gtcgtgccc aa 1212
164 <210> SEQ ID NO: 8
165 <211> LENGTH: 1647
166 <212> TYPE: DNA
167 <213> ORGANISM: Escherichia coli

```

RAW SEQUENCE LISTING
 PATENT APPLICATION US/09/701,132
 DATE 01/24/2001
 TIME 11:14:29

Input Set A:\P30384.app
 Output Set N:\CRF3\01232001\I701132.raw

```

169 <400> SEQUENCE: 8
170 atggcaaaag taatataac caacagcctc tgcctgatac ctcaaaaataa tatcaaacag 60
171 aaccagctctg cgtctgctgac ttctatcgag cgtctgtctt ctggtctgca tatcaaacag 120
172 gcaaaaggatg aaccccgagg tcaggcgatt gctaacggtt ttacttctaa catcaaacgg 180
173 ctgaactcag ctgcacgtaa ccccaacgac ggtatttctg ttgcgcagac caacgaaggc 240
174 gcctctgctc aaattaaaca caacttacag cgtattcctg aactgacggt tcaggcttct 300
175 acccgtaac aactcattc ggtctgac tcatttcagg acgaataca atccctctc 360
176 gacgaatttg acccgctac cgttcagacc cagttcaacg ggtgaacgt actggcaaaa 420
177 gacgcttcca tgaataatca ggttgggtcg aatgacggcc agctatcac tatcgactg 480
178 acaaaatctg actctgatac gctgggctcg aatgggttla atgtgaacgg caaaggggaa 540
179 acggttaata cgcgcagacc cctgaagat atgtctggtt tcaagctctg ggcggcaca 600
180 ggcgggaactg ttggtgtaac tcaatatac tcaaatctg ctgtaacaa tgcgttagat 660
181 attctaacat cgtctactg cgcagatgga aataaagata caactagcgc cgarcttgg 720
182 ttggttctac cgcgcgtctg tctaacctat aaccacaa aaagactaa tctatattc 780
183 ggcgtctctg atgataatc cgcgtctaac ctggtctgtt tctcaatcc tcaagccgga 840
184 gatacgaact aactcagct tacaattggt ggcgaagatc aagatgtaa catcgataa 900
185 tccggttaat taactgctcg tcatgctggt caactactt atctggtgc taccggtaac 960
186 ttacttaaaa ataatgctg tctgataaca caagctact tggctaaaact tgcctactct 1020
187 actggtgcta agcgcgcgac catcaaacct gataaaggaa catcaacaa tgcgggtaca 1080
188 gcatctgctg ggcctacat gtcctattgt acaatacat ttgcgaatgc agtaaaaat 1140
189 gacacttata ctgcactgt aggtgctaac acttatgctg taacaaacag ttctgctgct 1200
190 gacagctccg ctlatatgag caatggggtt ctcaagtata ctccgcacac ttactatgca 1260
191 caagctgctg gaagatcac aactactgag gatcgcgctg cgggttaact ggtctacaaa 1320
192 ggttccgctg gtaagttaac aacggatag actagcaaac cagaatcaac atcagatccg 1380
193 ctgcagctc ttgacgacg ctacagccag atgcacaaat cccgctctc cctgggtgct 1440
194 gtcgaacac ctctggttcc cgcagtgacc aacctgaaca aacacttac caactgtct 1500
195 gaaaggaact cgcgtatcca ggcgcgcac tatgcgacg aagtgtccaa catctcgaa 1560
196 ggcgaagatta tccagcaggc cgttaactcc gtcgtgccc aaagtaacca ggttccgca 1620
197 caggttctgt ctctgctgca ggtttaa 1647
199 <210> SEQ ID NO: 9
200 <211> LENGTH 1758
201 <212> TYPE DNA
202 <213> ORGANISM Escherichia coli
204 <400> SEQUENCE: 9
205 atggcaaaag taatataac caacagcctc tgcctgatac ctcaaaaataa tatcaaacag 60
206 aaccagctctg cgtctgctgac ttctatcgag cgtctgtctt ctggtctgca tatcaaacag 120
207 gcaaaaggatg aaccccgagg tcaggcgatt gctaacggtt ttacttctaa catcaaacgg 180
208 ctgaactcag ctgcacgtaa ccccaacgac ggtatttctg ttgcgcagac caacgaaggc 240
209 gcctctgctc aaattaaaca caacttacag cgtattcctg aactgacggt tcaggcttct 300
210 accggaacta actctgattc ggalctgac tcaattcagg acgaataca atccctctc 360
211 gatgaatttg acccgctac cgcgcagacc cagttcaacg ggtgaacgt actggcaaaa 420
212 gacggttcca tgaataatca ggttgggtcg aatgacggcc aaactatcac tatcgactg 480
213 aagaaatctg attctgatac tctgggtctg aatgggttla atgtaatgg taaggtact 540
214 atttaacaca aagctgcaac gctaaagtat ttaactctg ctgacgcgaa gtttaaacac 600
215 accagcggtc ttatgctct gaaaacgcaa ataacctgt taactacgca tgcgtgact 660
216 gataaattag ggaatggcga taagtcacac ctgacggcg tgaattater ttcaaacct 720
217 aactctggtg atttctact cagcaatct actgctgata cgggtgtaga cgcgcggcg 780
218 caagctactg attcagctaa aaaaactgat ggtttagct ccaaccttca tgcgtgctg 840
219 gtttaaatctg ttaactgttc ttaacacaca aaagatggtc ctgttctctt cgaacaggat 900

```

RAW SEQUENCE LISTING
 PATENT APPLICATION US/09/701,132

DATE: 01/23/2001
 TIME: 11.14.29

Input Set: A:\P30384.app
 Output Set: N:\CRF3\01232001\I701132.raw

```

220 tcagcagjta atatacccat cgggtggaagc caggcatatc taagcagatc agacacctta 960
221 acgactaaca acgctggtag ccgagctaaa gctgatatga aagcgcctgc taagcccgcg 1020
222 aacgaaggtg qtagcagtag ttctctgaca ttcaatggca ctgaatatat talcgcaaaa 1080
223 gcaactcttg cgaacacctc tccagttagt ccgttaacac ctgctgggat tacttatcag 1140
224 gctacagtga qtaaaagatg agtattgagc gaacccaaag cggctgcgcg gacatcttca 1200
225 attaccltta attcgggtgt actgaacaaa acatttgggt ttaccgcggg tgaatccagt 1260
226 gatgctacga agtcttatgt ggaatgataa ggtggtatta ctaccgttgc cgaactatac 1320
227 gtctcttaca ccgttaacaa ggaacacgac tctgtagctg ttgcgcggta tgcttcagcg 1380
228 actgatacaa ataaagatta tgcctcagca attggtactg ctgtaaatgt gaactcccg 1440
229 ggtlaaacca ctactgaagc taccagtgtt ggtcttgcaa cgaacaaacc gcttccatcc 1500
230 ctgagcagcg ctatcagctc caatgacaaa ttccgttctt cctctgggtg talcccaaac 1560
231 cgtctggatt ccgagctcac caactgaaac aacacacata ctacactgtc tgaagctgac 1620
232 tcccgatttc aggaacgcga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
233 atccagcaag ccgttaactc cgtgctggca aaagccaaac aggtacgcga qcaggttcta 1740
234 tctctgctgc aggtttaa 1758
236 <210> SEQ ID NO 70
237 <211> LENGTH 1383
238 <212> TYPE DNA
239 <213> ORGANISM Escherichia coli
241 <400> SEQUENCE 10
242 aacaaatctc agctctctct tagctctgct attgaacgct tctctcttgg tctggtatc 60
243 aacagcgcaa agcagcagtc agcagctcag gcaactgcta accgtttatc qcaaaatatt 120
244 aaaggtctga cccaggtctc ccgttaacga aatgatqcta ttctgtttgc qcaagccact 180
245 gaaggttcgc tgaatgaatc taacaaacac ctgcagcgta ttctgtgaact ttctgttcaq 240
246 gcaactaacg gtaactaactc tgacagcgat ctltcttcta tccagctcga aattactcaa 300
247 cgtctggaag aaattgacgc tctatctgag caaacctcag ttaacggcgt gaaagtcttt 360
248 gttgaaacta atgaatgaa aatttcaggt qttgctaatg atggtgaaac catcactatc 420
249 aatctgcaaa aaattgctgc gaaactcttc ggcctggaac qtttctaatc ccatgcgcgc 480
250 caaaaagcaa caggcagtgat cctgatttct aaatttcaag cgaacaggtac tgataattat 540
251 gatgttcgag gtaaaactta taccgtgaut gtcgagagcg gcgcggttaa gaatgatgct 600
252 aatlaaagtg ttcttctgaa ccagctgctt ggaatgcctg cgaacagtag tgaatctaaa 660
253 gtatccggtg aaagtattga tgaacacgaa ctacgcgaac ttgcaataaa attagctgac 720
254 aaaggctcca ttgatacaaa ggcattaca ttactaaca acactggcgc agagcttgat 780
255 gctaatggca aaggtgtttt gaccgcaaat attgatggtc aagatcttca atttactatt 840
256 gacagtatg caaccacggg tgcgcgcgca caactaaca cagacacagc tgttcaaaa 900
257 accagtgcgc gccagttcac caactaaaaa gtgaaaata aagcgcgaac actctctgct 960
258 ctggatctta atgcagccaa gaaacacagt agcaatttag ttgtaatagg cgcacactac 1020
259 aatgtcagcg taatgtgtaa aacgttaact gatactactc ctggtacccc taagtgtatg 1080
260 tatctgagca aatcagaagg tctgtgcgcg attctggtaa acgaagatgc agcaaaatcg 1140
261 ttgcaactta ccaccaaccc gctcgaaact atcgacaagg cattggctaa agttgacaat 1200
262 ctgcgttctg acccgggtgc agtacaacac cgtttctgac ctgccatcac caaccttgac 1260
263 aacaccgtaa caaacctctc ttctgcgcgt agcgttatcg aaagtctgca ctacgcgac 1320
264 gaagtgtcta acatctctcg tccgcagatc ctgcaacaag cgggtacctc tgtctcggcg 1380
265 cag 1383
267 <210> SEQ ID NO 11
268 <211> LENGTH 2013
269 <212> TYPE DNA
270 <213> ORGANISM Escherichia coli
272 <400> SEQUENCE 11

```

L:9 M:270 C. Current Application Number differs, Replaced Application Number
L:10 M:271 C. Current Filing Date differs, Replaced Current Filing Date

[illegible]

09/701132

522 Pm'd Pm'd Pm'd 21 NOV 2000

SEQUENCE LISTING

<110> THE UNIVERSITY OF SYDNEY

<120> ANTIGENS AND THEIR DETECTION

<130> REEVES

<140>

<141>

<160> 68

<170> PatentIn Ver. 2.0

<210> 1

<211> 1773

<212> DNA

<213> Escherichia coli

<400> 1

```

atgcgacgta tagaacgaat accgggggta tcggcgtaag cggggcacaag tttagcattt 60
attttttggc ttaatgacac gaacagcaac gaggaagggg agtatttcga ccgctagaaa 120
aaaattctaa aggttgtgag tgaccagacg ataacagggg tgacggcgac gaagccgaag 180
ggtggaagcc caatacttaa accgtagact tgaaaacagg aaaatgaatc atggcacaag 240
tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag aaccagtctg 300
cgctgtcgac ttctatcgag cgctctctt ctggtctgag cattaacagc gctaaagatg 360
acgctgaggg ocaagcgatt gctaaccgct tcacttctaa catcaaaggt ctgactcagg 420
ccgcacgtaa cgccaacgac ggtatttctc tggcgagac cactgaaggc gactgtctg 480
aaatcaacaa caacttgacg cgtgttcgtg aactgaccgt tcaggccact accggtacta 540
actctgattc tgacctgtct tcaatacagg acgaaatcaa atcccgctct gatgaaattg 600
accgcgtatc cggtcagact cagttcaacg gcgttaatgt tctttccaaa gatggttcaa 660
tgaaaattca ggttggtgag aatgatggtc aaactatctc catcgatctg aagaaaattg 720
attcttcaac tttggggctg aatggcttct cagtttctaa aaactctctt aatgtcagca 780
atgctatcac atctatccc caagccgcta gcaatgaacc tgttgatggt aacttcgggt 840
atactgatga gtctgcagca atcgagcca aattgggggt ttccgatacg tcaagcctgt 900
cgctgcacaa catccttgat aaagatggta aggcaacagc tgattatggt gttcagtcag 960
gtaaagactt ctatgctgct tctgttaatg ccgcttcagg taaagtaacc ttaaaccacca 1020
ttgatgttac ttatgatgat tatgcgaacg gtgttgacga tgccaagcaa acaggtcagc 1080
tgatcaaagt ttccagcagat aaagacggcg cagctcaagg ttttgtcaca cttcaaggca 1140
aaaactattc tgctgggtgag gcgagacaca ttcttaagaa tggagcaaca gctcttaagt 1200
taactgatct gaatttaagt gatgttactg atactaatgg taaggtaacc acaactgcga 1260
ctgagcaatt tgaaggtgct tcaactgagg atccgctggc gcttctggat aaagctattg 1320
catcagtcga caaattccgg tcttctctag gtgcccgtgca gaaccgtctc gattccgcta 1380
tcaccaacct gaacaacacc accaccaacc tgtctgaagc gcagtcacct attcaggacg 1440
ccgactatgc gaccgaagtg tccaacatgt cgaagcgca gatcatccag caggcaggta 1500
actccgtgct gtctaaagcg aaccagggtac cgcagcaagt tctgtcactg ttacaaggct 1560
aatggcctta acctgcctga cccgccacc ggcggggttt tttctgtccg caatttaccg 1620
ataaccccc aataacccct catttcaccc actaatcgct cgattaaaaa ccctgcagaa 1680
acggataatc atgcccataa ctcatataac gcagggtgtt ttatcgtgaa ttcactctat 1740
accgctgaag gtgtaatgga taaacactcg ctg

```

<210> 2

<211> 500

<212> DNA

<213> Escherichia coli

<400> 2

```

aacagcctct cgctgatcac tcagaacaac atcaacaaaa accagtcttc aatgtctact 60
gccattgagc gtctgtcttc cggctctgct atcaacagcg caaaagatga cgctgctggc 120
caggcgattg ccaaccgctt cactctaac atcaaaagtc tgactcaggc agctcgtaac 180
gccaacgacg gtatctccgt tgcacagacc actgaaggcg cactgtctga aatcaacaac 240
aacctgcagc gtatccgtga gctgactggt cagtcttcta cgggtactaa ctctgaatcc 300
gatctgaact caatccagga cgaaattaaa tcccgtctgg acgaaattga ccgctatcc 360
ggtcagaccc agttcaacgg cgtgaacgtg ctggcaaaag acggctccat gaaaattcag 420
gttggcgcg aagatggtga aaccatcacc atcgacctga aaaaaattga ctcttctact 480
ttaaacctga ctgggtttta

```

<210> 3
 <211> 500
 <212> DNA
 <213> Escherichia coli

<400> 3
 ctcagtatgc tgtcaccggc agtacagggt ccgtaactta cgatccagat acagatcctg 60
 ccgcgactgg tgatattggt tctgcttatg ttgatgatgc aggtacattg acaactgatg 120
 caaacaaaac tgtaaaatat tatgcccaca ctaatggtag cgtcacgaac gacagtgggt 180
 cagctattta cgcaactgaa gcgggcaaat tgactactga agcgtctaca gctgctgaaa 240
 ctaccgctaa cccactgaaa gccctggacg atgcaatcag ccagatcgac aaattccgtt 300
 cttctctggg tgctgtacag aaccgtctgg attctgcggt aaccaacctg aacaacacca 360
 ccaccaacct gtctgaagcg cagtcccgta ttcaggacgc cgactatgcg accgaagtgt 420
 caaatatgtc taaagcgagc atcatccagc aggcgggtaa ctccgtgttg gctaaagcta 480
 accaggttcc tcagcaggtt 500

<210> 4
 <211> 399
 <212> DNA
 <213> Escherichia coli

<400> 4
 agcctgtgcg tgttgacca gaataacctg aacaaatctc agtcttctct gagctccgcc 60
 attgagcgct tctcttctgg cctgcgtatt aacagtgcga aagatgacgc agcaggtcag 120
 gcgattgcta accgttttac agcaaatatt aaaggctctga ctcaggcttc ccgtaacgcg 180
 aatgatggta tttctgttgc gcagaccact gaaggcgcgc tgaatgaaat taacaacaac 240
 ctgcagcggt tacgtgaact gactgttcag gcaactaacg gtactaactc tgacagcgat 300
 ctttcttcta tccaggctga aattactcaa cgtctggaag aaattgaccg tgtatctgag 360
 caaactcagt ttaacggcgt gaaagtcctt gctgaaaat 399

<210> 5
 <211> 417
 <212> DNA
 <213> Escherichia coli

<400> 5
 gcacgttagt tgttaacggg gcaacttacg atgttagtgc agatggtaaa acgataacgg 60
 agactgcttc tggtaacaat aaagtcatgt atctgagcaa atcagaagggt ggtagcccca 120
 ttctggtaaa cgaagatgca gcaaaatcgt tgcaatctac caccaacccg ctcgaaacta 180
 tcgacaaagc attggctaaa gttgacaatc tgcgttctga cctcggtgca gtacaaaacc 240
 gtttcgactc tgctatcacc aaccttggca acaccgtaaa caacctgtct tctgccccta 300
 gccgtatcga agatgctgac tacgcgaccg aagtgtctaa catgtctcgt gcgcagatcc 360
 tgcaacaagc ggggtacctc gttctggcgc aggctaacca gaccacgcag aacgtac 417

<210> 6
 <211> 950
 <212> DNA
 <213> Escherichia coli

<400> 6
 aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgtatt 60
 aacagcgcta aagatgacgc cgcgggccag gcgattgcta accgctttac ttctaacatc 120
 aaaggctctga ctcaggccgc acgtaacgcc aacgacggta tttctctggc gcagacggct 180
 gaaggcgcgc tgtcagagat taacaacaac ttgcagcgta ttcgtgaact gaccgttcag 240
 gccctctacc gcacgaactc tgattccgac ctgtcttcta ttcaggacga aatcaaatcc 300
 cgtcttgatg aaattgaccg tgtatctggt cagacccagt tcaacgggtg gaacgtgctg 360
 tcgaaaaacg attcgaatgaa gattcagatt ggtgccaatg ataaccagac gatcagcatt 420
 ggcttgcaac aaatcgacag taccactttg aatctgaaag gatttaccgt gtccggcatg 480
 gcggatttca gcgcggcgaa actgacggct gctgatggta cagcaattgc tgctgcggat 540
 gtcaaggatg ctgggggtaa acaagtcaat ttactgtctt acactgacac cgcgtctaac 600
 agtactaaat atgcggtcgt tgattctgca accggtaaat acatggaagc cactgtagtc 660
 attaccggta cggcggcggc ggtaactggt ggtgcagcgg aagtggcggg agccgctaca 720
 gccgatccgt taaaagcact ggatgccgca atcgctaaag tcgacaaatt ccgctcctcc 780
 ctcggtgcgg ttcaaaaccg tctggattct gcggtcacca acctgaacaa caccaccacc 840
 aacctgtctg aagcgcagtc ccgtattcag gacgccgact atgcgaccga agtgtccaac 900
 atgtcgaaa ggcagattat ccagcaggcg ggcaactccg tgctgtctaa 950

<210> 7
 <211> 1212
 <212> DNA
 <213> Escherichia coli

<400> 7
 aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgtatt 60
 aacagcgcta aagatgacgc cgcgggccag gcgattgcta accgcttcac ttctaacatc 120
 aaaggctctga ctcaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180
 gaaggcgcgc tgtctgaaat caacaacaac ttgcagcgctg tgcgtgagtt gaccgttcag 240
 gcgacgaccg ggactaactc tgattctgac ctgtcttcta ttcaggacga aatcaaatcc 300
 cgtctggatg aaattgatcg cgtttccggg cagacccagt tcaacggcgt gaatgtgctg 360
 gcgaaagatg gttcgatgaa gattcagggt ggcgcaatg atgggcagac tattagcatt 420
 gatttgacaga agattgactc ttctacatta ggactgaacg gtttctccgt ttcgggtcag 480
 tcacttaacg ttagtgattc cattactcaa attaccgggt ccgcccggac aaaacctgtt 540
 ggtgttgatt tcaactgctg tgcgaaagat ctgactactg cgacaggtaa aacagtcgat 600
 gtttctagcc tgacgttaca caacactctg gatgcgaaag gggctgctac atcacagttc 660
 gtcgttcaat ccggcaatga tttctactcc gcgtcgatta atcatacaga cggcaaaagt 720
 acgttgataa aagccgatgt cgaatacaca gacaccgata atggactaac gactgcggct 780
 actcagaaag atcaactgat taaagttgcc gctgactctg acggctcggc tgcgggatat 840
 gtaacattcc aaggtaaaaa ctacgctaca acggtttcaa cggcacttga tgataatact 900
 gcggcaaaag caacagataa taaagttggt gttgaattat caacagcaaa accgactgca 960
 cagttctcag gggcttcttc tgctgatcca ctggcacttt tagacaaagc tattgcacag 1020
 gttgatactt tccgctcctc cctcggtgcg gtgcaaaacc gtctggattc cgcagtaacc 1080
 aacctgaaca acaccaccac caacctgtct gaagcgcagt cccgtattca ggacgccgac 1140
 tatgtctacg aaggttccaa catgtcgaaa gcgcagatca tccagcaggc aggttaactcg 1200
 gtgctgtcca aa 1212

<210> 8
 <211> 1647
 <212> DNA
 <213> Escherichia coli

<400> 8
 atggcacaaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagtcctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
 gcgaaggatg acgcccgggg tcaggcgatt gctaaccgtt ttactttctaa cattaaaggc 180
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
 gcgctgtccg aaattaacaa caacttacag cgtattcgtg aactgacggg tcaggcttct 300
 accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
 gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
 gacggttcga tgaaaattca ggttggtgag aatgacggcc agactatcac tattgatctg 480
 aagaaaattg actctgatac gctggggctg aatgggttta atgtgaacgg caaaggggaa 540
 acggctaata cggcagcaac cctgaaagat atgtctggat tcacagctgc ggcggcacca 600
 gggggaaactg ttggtgtaac tcaatatact gacaaatcgg ctgtagcaag tagcgtagat 660
 attctaaatg ctggtgctgg cgcatatgga aataaagtta caactagcgc cgatgttggt 720
 tttggtacac cagccgctgc tgtaacctat acctacaata aagacactaa ttcatattcc 780
 gccgcttctg atgatatttc cagcgtaac ctggctgctt tcctcaatcc tcaggccgga 840
 gatacgacta aagctacagt tacaattggt ggcaaaagat aagatgtaaa catcgataaa 900
 tccggttaatt taactgctgc tgatgatggc gcagtacttt atatggatgc taccggtaac 960
 ttaactaaaa ataatgctgg tggtgatata caagctactt tggctaaact tgctactgct 1020
 actggtgcta aagccgcgac catccaaact gataaaggaa cattcaccag tgacggtaac 1080
 gcgtttgatg gtgcatcaat gtccattgat accaatacat ttgcaaatgc agtaaaaaat 1140
 gacacttata ctgccactgt aggtgctaag acttatagcg taacaacagg ttctgctgct 1200
 gcagacaccg cttatatgag caatgggggt ctcatgata ctccgccaac ttactatgca 1260
 caagctgatg gaagtatcac aactactgag gatgcggctg ccggtaaaact ggtctacaaa 1320
 ggttccgatg gtaagttaac aacggatacg actagcaaa cagaatcaac atcagatccg 1380
 ctggcagctc ttgacgacgc tatcagccag atcgacaaat tccgctcctc cctgggtgctg 1440
 gtgcaaaacc gtctggattc cgcagtgaac aacctgaaca acaccactac caacctgtct 1500
 gaagcgcagt cccgtattca ggacgccgac tatgcgaccg aagtggtccaa catgtcgaaa 1560
 gcgcagatta tccagcaggc ggtaactcc gtgctggcaa aagctaacca ggttccgcag 1620
 caggttctgt ctctgctgca ggggttaa 1647

<210> 9
 <211> 1758
 <212> DNA
 <213> Escherichia coli

<400> 9
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagctcg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
 gcgaaggatg acgccgcggg tcaggcgatt gctaaccggt ttacttctaa cattaagggc 180
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac caccgaaggc 240
 gcgctgtctg aaatcaacaa caacttacag cgtatccgtg agctgacggg tcaggcttct 300
 accggaacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctt 360
 gatgaaattg accgcgtatc cggccagacc cagtccaacg gcgtgaacgt actggcaaaa 420
 gacggttcga tgaaaattca ggttggtgcg aatgacgggt aaactatcac tatcgacctg 480
 aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaaatgg taaaggtaact 540
 attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaaccac 600
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
 gataaattag ggaatggcga taaagtcaac gttggcggcg tagattatac ttacaacgct 720
 aaatctggtg attttactac caccaaactc actgctggta cgggtgtaga cgccgcggcg 780
 caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
 ggtaaatctg ttaatgggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
 tcagcaggta atatcaccat cgggtggaagc caggcatacg tagacgatgc aggcaacttg 960
 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgctgct taaagccgcg 1020
 agcgaaggta gtgacgggtc ttctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
 gcaactctcg cgacaacctc tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
 gttacagtga gtaaaagtgt agtattgagc gaaaccaaag cggctgcccg gacatcttca 1200
 attaccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
 gatgctgcga agtcttatgt ggatgataaa ggtggtatta ctaacggtgc cgactataca 1320
 gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
 actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440
 ggtaaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
 ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
 cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgag 1620
 tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
 atccagcagg ccggtaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
 tctctgctgc aggggttaa 1758

<210> 10
 <211> 1383
 <212> DNA
 <213> Escherichia coli

<400> 10
 aacaaatctc agtcttctct tagctctgct attgagcgtc tgtcttctgg tctgcgtatt 60
 aacagcgcaa aagacgatgc agcaggctcag gcgattgcta accgttttac ggcaaatatt 120
 aaaggctctga cccaggcttc ccgtaacgca aatgatggta tttctgttgc gcagaccact 180
 gaagggtgcg tgaatgaaat taacaacaac ctgcagcgta ttctgtaact ttctgttcag 240
 gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggctga aattactcaa 300
 cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360
 gctgaaaata atgaaatgaa aattcaggtt ggtgctaata atggtgaaac catcactatc 420
 aatctggcaa aaattgatgc gaaaactctc ggcctggacg gttttaatat cgatggcgcg 480
 cagaaagcaa caggcagtga cctgatttct aaattttaaag cgacaggtag tgataattat 540
 gatgttggcg gtaaaactta taccgtgaat gtggagagcg gcgcggttaa gaatgatgct 600
 aataaagatg tttttgtaag cgcagctgat ggatcgctga cgaccagtag tgatactaaa 660
 gtatccggtg aaagtattga tgcaacagaa ctagcgaac ttgcaataaa attagctgac 720
 aaaggctcca ttgaatacaa gggcattaca ttactaaca acactggcgc agagcttgat 780
 gctaattgga aagggtgttt gaccgcaaat attgatggtc aagatgttca atttactatt 840
 gacagtaatg caccacggg tgccggcgca acaataacta cagacacagc tgtttacaaa 900
 aacagtgcgg gccagttcac cactacaaaa gtggaaaata aagccgcaac actctctgat 960
 ctggatctta atgcagccaa gaaaacaggt agcactttag ttgtaaatgg cgccacctac 1020
 aatgtcagcg cagatggtaa aacggtaact gatactactc ctgggtgccc taaagtgatg 1080
 tatctgagca aatcagaagg tggtagcccg attctggtaa acgaagatgc agcaaaatcg 1140
 ttgcaatcta ccaccaaccc gctcgaaact atcgacaagg cattggctaa agttgacaat 1200
 ctgcgttctg acctcggtgc agtacaaaac cgtttctgact ctgccatcac caaccttggc 1260
 aacaccgtaa acaacctgtc ttctgcccgt agccgtatcg aagatgctga ctacgcgacc 1320
 gaagtgtcta acatgtctcg tgcgcagatc ctgcaacaag cgggtacctc tgttctggcg 1380
 cag 1383

<210> 11
 <211> 2013
 <212> DNA
 <213> Escherichia coli

<400> 11
 atggcacaag tcattaatac caacagcctc tgcgtgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggccttgcg tattaacagc 120
 gcgaaggatg acgccgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttccg ttgcacagac cactgaaggc 240
 gcgctgtccg aaattaacaa caacttacag cgtattcgtg aactgacggg tcaggcttct 300
 accgggacta actccgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
 gacgaattg accgcgtatc cgccagacc cagttcaacg gcgtgaacgt gctgtccaaa 420
 gatggctoga tgaaaattca ggtcggcgcg aacgatggcg aaacgattac tattgatctg 480
 aagaaaattg actctgatac gctgaatctg gctggtttta acgttaacgg taaaggttct 540
 gtagcgaata cagctgcgac aagcgacgat ttaaaactgg ctgggtttcac taagggcacc 600
 acagatacca atggcgtgac cgcgtataca aacacaatta gtaatgacaa agccaaagct 660
 tccgatctgt tagctaatat caccgatgga tcagtgatca ctgggggagg ggcaaacgct 720
 tttggcgtgg ctgcaaagaa tggttacacc tatgatgcag caagtaaatc ttatagtttt 780
 gctgcagatg gtgccgattc agcgaagacg ttaagcatca ttaatccaaa caccggtgat 840
 tcgtcgcagg cgacagtgcg tattggtggg aaagagcaga aagttaatat ttcccaggat 900
 ggaaaaatta ctgcgcgaga tgataatgcg acgctgtatt tagataaaca gggaaacttg 960
 acaaaaacga atgcaggtaa cgataccgca gcgacttggg atggtttaat ttccaacagc 1020
 gattctaccg gtgcggttcc agttgggggt gcaactacaa ttacaattac ttctggtaca 1080
 gcttccggaa tgtctgttca gtccgcagga gcaggaattc agacctcaac aaattctcag 1140
 attcttgtag gtggtgcatt tgcggctaag gtaagtattg agggaggcgc tgctacagac 1200
 attttgtag caagtaaatg aaacataaca gcggctgatg gtagtgact ttatcttgat 1260
 ggcactactg gtggattcac tacaacggct ggaggaaata cagctgcttc gttagataat 1320
 ttaattgcta acagtaagga tgctacctta accgtaactt caggtaaccg ccagaacact 1380
 gtttatagca caacaggaag tggcgctcag ttcaccagtt tagcaaaagt agacacagtc 1440
 aatgtcacca acgcacatgt cagtgcgaa ggtatggcaa atctgacaaa aagcaatttt 1500
 accattgata tgggcggtac aggtacagta acttacacag tttccaatgg ggatgtgaaa 1560
 gctgctgcaa atgctgatgt ttatgtcgaa gatggtgcac tttcagccaa tgctacaaa 1620
 gatgtaacct actttgaaca aaaaaatggg gctattacca acagcaccgg tggtagcatc 1680
 tatgaaacag ctgatggtaa gttaacaaca gaagctacta ctgcatccag ttccaccgcc 1740
 gatcccctga aagctctgga cgaagccatc agctccatcg acaaattccg ctctccctc 1800
 ggtgcggtgc aaaaccgtct ggattccgcg gtcaccaacc tgaacaacac cactaccaac 1860
 ctgtccgaag cgcagtcctg tattcaggac gccgactatg cgaccgaagt gtccaacatg 1920
 tcgaaagcgc agatcatcca gcagccggg aactccgtgc tggcaaaagc taaccaggta 1980
 ccgcagcagg ttctgtctct gctgcagggt taa 2013

<210> 12
 <211> 1263
 <212> DNA
 <213> Escherichia coli

<400> 12
 atggcacaag tcattaatac caacagcctc tgcgtgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggccttgcg tattaacagc 120
 gcgaaggatg acgccgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
 gcgctgtccg aaattaacaa caacttacag cgtgtgcgtg agctgactgt tcaggcgacc 300
 accgggacta actctgagtc tgacctgtct tctatccagg acgaaatcaa atctcgctg 360
 gaagagattg atcgtgtttc aagtcagact caatttaacg gcgtgaatgt tttggctaaa 420
 gatgggaaaa tgaacattca ggttggggca aatgatggac agactatcac tattgatctg 480
 aaaaagatcg attcatctac actaaacctc tccagttttg atgtacaaa cttgggcacc 540
 agtgttaaag atggggccac catcaataag caagtggcag taggtgctgg cgactttaaa 600
 gataaagctt caggatcgtt aggtacccta aaattagttg agaaagacgg taagtactat 660
 gtaaatgaca ctaaaagtag taagtactac gatgccgaag tagatactag taagggtaaa 720
 attaaacttca actctacaaa tgaaagtgga actactccta ctgcagcgac ggaagtaact 780
 actgttggcc gcgatgtaaa attggatgct tctgcactta aagccaacca atcgcttgc 840
 gtgtataaag ataaaagcgg caatgatgct tatatcatc agaccaaaga tgtacaact 900
 aatcaatcaa ctttcaatgc cgctaatac agtgatgctg gtgttttata tattggtgca 960

tctacaaccg	cgccaagcaa	tttaacagct	aacccgctta	aggctcttga	tgatgcaatt	1020
gcattctgtt	ataaattccg	ctcttctctc	ggtgccgttc	agaaccgtct	ggattctgcc	1080
attgccaacc	tgaacaacac	cactaccaac	ctgtctgaag	cgagtcctcg	tattcaggac	1140
gctgactatg	cgaccgaagt	gtccaacatg	tcgaaagcgc	agattatcca	gcaggccggg	1200
aactccgtgc	tgccaaaagc	caaccaggta	cgcagcaggt	ttctgtctct	gctgcagggt	1260
taa						1263

<210> 13
 <211> 1368
 <212> DNA
 <213> Escherichia coli

<400> 13						
aacaaatctc	agtcttctct	gagctccgcc	attgaacgtc	tctcttcttg	cctgcgtatt	60
aacagtgcga	aagatgacgc	agcagggtcag	gcgattgcta	accgtttttac	agcaaataatt	120
aaagggtctga	ctcaggcttc	ccgtaacgcg	aatgatggta	tttctgttgc	gcagaccact	180
gaagggtgcgc	tgaatgaaat	taacaacaac	ctgcagcgtg	tacgtgaact	gactgttcag	240
gcaactaacg	gtactaactc	tgacagcgat	ctttcttcta	tccagggtga	aattactcaa	300
cgtctggaag	aaattgaccg	tgtatctgag	caaaactcagt	ttaacggcgt	gaaagtcctt	360
gctgaaaaata	atgaaatgaa	aattcagggt	ggtgctaattg	atgggtgaaac	catcactatc	420
aatctgggcaa	aaattgatgc	gaaaactctc	ggcctggacg	gttttaatat	cgatggcgcg	480
cagaaagcaa	ctggcagtg	cctgatttct	aaattttaaag	cgacagggtac	tgataactat	540
gatgttggtg	gtgatgctta	tactgttaac	gtagatagcg	gagctgttaa	agatactaca	600
gggaatgata	tttttggttag	tgacagcagat	ggttcactga	caactaaatc	tgacacaaac	660
atagctggta	cagggattga	tgctacagca	ctcgcagcag	cggctaagaa	taaagcacag	720
aatgataaat	tcacgtttta	tggaagttaa	ttcacaacaa	caactgcagc	ggatggcaat	780
gggaatgggt	tatatctctg	agaaattgat	ggtaagtcag	tgacattttac	tgtgacagat	840
gctgacaaaa	aagcttcttt	gattacagag	gagacagttt	acaaaaatag	cgtgtggcctt	900
tatacgacaa	ccaaagttga	taacaaggct	gccacacttt	ccgatcttga	tctcaatgca	960
gctaagaaaa	caggaagcac	gttagttggt	aacgggtgcaa	cttacgatgt	tagtgacgat	1020
ggtaaaaacga	taacggagac	tgcttctggt	aacaataaag	tcattgtatct	gagcaaatca	1080
gaagggtgga	gcccgaattct	ggtaaacgaa	gatgcagcaa	aatcgttgca	atctaccacc	1140
aacccgctcg	aaactatcga	caaagcattg	gctaaagttg	acaatctgcg	ttctgacctc	1200
ggtgcagtac	aaaaccgttt	cgactctgct	atcaccaacc	ttggcaaac	cgtaaacac	1260
ctgtcttctg	cccgtagccg	tatcgaagat	ctgactacg	cgaccgaagt	gtctaacaatg	1320
tctcgtgcgc	agatcctgca	acaagcgggt	acctctgttc	tggcgcag		1368

<210> 14
 <211> 1788
 <212> DNA
 <213> Escherichia coli

<400> 14						
atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagtctg	cgtgtgcgag	ttctatcgag	cgtctgtctt	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgcagcggg	tcaggcgatt	gctaaccgtt	tcacctctaa	cattaagggc	180
ctgactcagg	cggcccgtaa	cgccaacgac	ggtatctccg	ttgcgcagac	caccgaaggc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgtatccgtg	aactgacggg	tcaggcttct	300
accgggacta	actccgattc	ggatctggac	tcatttcagg	acgaaatcaa	atcccgtctg	360
gacgaaattg	accgcgtatc	tgccagacac	cagttcaacg	gcgtgaacgt	actggcgaaa	420
gacgggtcaa	tgaaaattca	ggttggtgcg	aatgacggcc	agactatcac	gattgatctg	480
aagaaaattg	actcagatac	gctggggctg	aatgggttta	acgtgaatgg	ttccggtagc	540
atagccaata	aagcggcgac	cattagcgac	ctgacagcag	cgaaaatgga	tgctgcaact	600
aatactataa	ctacaacaaa	taatgcgctg	actgcataca	aggcgcttga	tcaactgaaa	660
gatggtgaca	ctgttactat	caaagcagat	gctgctcaaa	ctgccacggg	ttatacatat	720
aatgcatcag	ctggtaactt	ctcattcagt	aatgtatcga	ataataactc	agcaaaagca	780
ggtgatgtag	cagctagcct	tctcccgcgc	gctgggcaaa	ctgctagtgg	tgttttataa	840
gcagcaagcg	gtgaagtga	ctttgatggt	gatgcgaatg	gtaaaatcac	aatcggagga	900
cagaaagcat	atttaactag	tgatggtaac	tttaactaaa	acgatgctgg	tggtgcgact	960
gcggctacgc	ttgatggttt	attcaagaaa	gctgggtgatg	gtcaatcaat	cgggttttaag	1020
aagactgcat	cagtcacgat	ggggggaaca	acttataact	ttaaaacggg	tgctgatgct	1080
gatgctgcaa	ctgctaaccg	aggggtatcg	ttcactgata	cagctagcaa	agaaaccgtt	1140
ttaaataaag	tggttacagc	taaacaaggc	aaagcagttg	cagctgacgg	tgatacatcc	1200
gcaacaatta	cctataaatc	tgcggttcag	acgtatcagg	ctgtatttgc	cgcagggtgac	1260
ggtactgcta	gcgcaaaata	tgccgataaa	gctgacgttt	ctaattgcaac	agcaacatac	1320

```

actgatgctg atggtgaaat gactacaatt ggttcataca ccacgaagta ttcaatcgat 1380
gctaacaacg gcaaggtaac tgttgattct ggaactggta cgggtaaata tgcgcgaaa 1440
gtaggggctg aagtatatgt tagtgctaatt ggtactttaa caacagatgc aactagcgaa 1500
ggcacagtaa caaaagatcc actgaaagct ctggatgaag ctatcagctc catcgacaaa 1560
ttccgttctt cctgggtgc tatccagaac cgtctggatt ccgcagtcac caacctgaac 1620
aacaccacta ccaacctgtc cgaagcgag tcccgtattc aggacgccga ctatgcgacc 1680
gaagtgtcca acatgtcgaa agcgcagatc attcagcagg ccggtaactc cgtgctggca 1740
aaagccaacc aggtaccgca gcaggttctg tctctgctgc aggggttaa 1788

```

```

<210> 15
<211> 1653
<212> DNA
<213> Escherichia coli

```

```

<400> 15
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccggt ttactttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttccg ttgcgcagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg agctgacggt tcaggcttct 300
accgggacta actccgattc tgacctggac tccatccagg acgaaatcaa gtctcgtctg 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacggttcga tgaaaattca ggttggtgcg aatgacggcc agactatcac gattgatctg 480
aagaaaattg actcagatac gctggggctg agtgggttta atgtgaatgg tggcggggct 540
gttgctaaca ctgctgcac taaagctgac ttggtagctg ctaatgcaac tgtggtaggc 600
aacaaatata ctgtgagtgc gggttacgat gctgctaaag cgtctgattt gctggctgga 660
gttagtgatg gtgatactgt tcaggcaacc attaataacg gcttcggaac ggcggctagt 720
gcaacgaatt acaagtatga cagtgcaggt aagtcttact cttttgatac cacaacggct 780
tcagctgccg atgttcagaa atatttgacc ccgggcgttg gtgataccgc taagggcact 840
attactatcg atggttctgc acaggatggt cagatcagca gtgatggtaa aattacgtca 900
agcaatggag ataaacttta cattgataca actgggcgct taacgaaaaa cggctttagt 960
gcttctttga ctgaggctag tctgtccaca cttgcagcca ataataccaa agcgacaacc 1020
attgacattg gcggtacctc tatctccttt accggttaata gtactacgcc gaacactatt 1080
acttattcag taacagggtgc aaaagttgat caggcagctt tcgataaagc tgtatcaacc 1140
tctggaacag atgttgattt cactaccgca ggttatagcg tcgacggcgc aactggcgct 1200
gtaacaaaag gtgttgctcc ggtttatatt gataacaacg gggcgttgac cacatctgat 1260
actgtagatt tttatctaca ggatgatggt tcagtgacta acggcagcgg taaggcagtt 1320
tataaagatg ctgacggtaa attgacgaca gatgctgaaa ctaaagctgc aaccaccgcc 1380
gatccccctg aagctctgga cgaagccatc agctccatcg acaaattccg ctctccctc 1440
ggtgcgggtg agaaccgtgc ggattcccg gtcaccaacc tgaaccaaac cactaccaac 1500
ctgtctgaag cgcagtcctc tattcaggac gctgactatg cgaccgaagt atccaacatg 1560
tcgaaagcgc agatcatcca gcaggccggt aactccgtgc tggcaaaagc taaccaggta 1620
ccacagcagg ttctgtctct gctgcagggt taa 1653

```

```

<210> 16
<211> 1689
<212> DNA
<213> Escherichia coli

```

```

<400> 16
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccggt ttactttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtgtgcgtg aactgaccgt tcaggcaacc 300
accggtagca actccagctc tgacctggac tctatccagg acgaaattaa atcccgtctg 360
gacgaaattg atcgcgtatc cggtcagacc cagttcaacg gcgtgaacgt gctggcaaaa 420
gacggttcga tgaaaattca ggttggcgcg aacgatggcc agaccatcac tatcgacctg 480
aagaagattg actcttctac cttgaacctg acaggtttta acgttaacgg ttctggttct 540
gtggcgata ctgcagcaac taaagctgat ttaaccgctg ctcaactctc tgcaccgggt 600
gcagcagacg caaatggtac agttacttat actgtcagtg ctggttataa agaatccact 660
gctgcagatg ttattgctag catcaaagac ggcagtgctc cgacttctgc aattactgca 720
accattaata atggcttcgg tgattccagt gcgctgactt ccaatgacta tacttatgac 780
ccagcaaaaag cgcacttcac ttacgacgta gcttcaagcg gcttcaagcg ccaataatac tgctgcccg 840
gttcagtcct tcctgacgcc gaaagcagggt gataccgcaa atctgaaagt aaccgttggt 900

```

```

acgacatcgg ttgatgtcgt tctggccagt gatggtaaga ttacagcaaa agatggttct 960
gcattatata tcgacagtac aggtaacctg actcagaaca gtgctggctt gacctctgct 1020
aaactggcta ctctgactgg ccttcagggc tctgggtgtg cttcaaccat cactactgaa 1080
gatggcacta atattgatat tgctgctaac ggtaatatgt gtctgaccgg tgttcgtatc 1140
agtgtgtgatt ctctgcagtc agcgactaaa tctacgggct ttactgttgg tactggcgct 1200
acaggtctga ccgtaggtac tgatggtaaa gtgactatcg gcgggactac tgctcagtc 1260
tacaccagca aagatggttc cctgactact gataacacca ctaaactgta tctgcagaaa 1320
gatggctctg taaccaacgg ttcaggtaaa gcggtctatg tagaagcgga tggtgatttc 1380
actaccgacg ctgcaaccaaa agcgcgaacc accaccgatc cgctgaaagc cctggatgag 1440
gcaatcagcc agatcgataa gttccgttca tccctgggtg ctatccagaa ccgtctggat 1500
tccgcgggtc caacctgaa caacaccact accaacctgt ctgaagcgca gtcccgtatt 1560
caggacgccc actatgcgac cgaagtgtcc aacatgtcga aagcgcagat cattcagcag 1620
gccggtaact ccgtgctggc aaaagccaac caggtaccgc aacaggttct gtctctgctg 1680
cagggcctaa                                     1689

```

```

<210> 17
<211> 915
<212> DNA
<213> Escherichia coli

```

```

<400> 17
gcgctgtcga cttctatcga gcgcctctct tctggctctgc gtattaacag cgctaaagat 60
gacgctgcgg gccaggcgat tgctaaccgc ttcacttcta acatcaaagg tctgactcag 120
gccgcacgta acgccaacga cggattttct ctggcgca gaagctgaagg cgcgctgtca 180
gagattaaca acaacttgca gcgtattcgt gaactgaccg ttcaggcctc taccggcacg 240
aactctgatt ccgacctgtc ttctattcag gacgaaatca aatcccgtct tgatgaaatt 300
gaccgtgtat ctggtcagac ccagttcaac ggtgtgaacg tgctgtcgaa aaacgattcg 360
atgaagattc agattggtgc caatgataac cagacgatca gcattggctt gcaacaaatc 420
gacagtacca ctttgaatct gaaaggattt accgtgtccg gcattggcga tttcagcgcg 480
gcgaaactga cggctgctga tggtagacga attgctgctg cggatgtcaa ggatgctggg 540
ggtaaacaag tcaatttact gtcttacct gacaccgct ctaacagtac taaatatgctg 600
gtcgttgatt ctgcaaccgg taaatacatg gcagccactg tagtcattac cagtacggcg 660
gcggcggtaa ctgttggtgc aacggaagtg gcgggagccg ctacagccga accgttaaaa 720
gcactggatg ccgcaatcgc taaagtgcac aaattccgct cctccctcgg tgcggttcaa 780
aaccgtctgg attctgcggt caccaacctg aacaacacca ccaccaacct gtctgaagcg 840
cagtcgccgt ttcaggacgc cgactatgct accgaagtgt ccaacatgtc gaaagcgag 900
attatccagc agggcg                                     915

```

```

<210> 18
<211> 1665
<212> DNA
<213> Escherichia coli

```

```

<400> 18
atggcacaaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccggt ttacttctaa tattaagggc 180
ctgactcagg ctgcacgtaa cgccaatgac ggtattttct ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggg tcaggccact 300
acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctg 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt gctgtccaaa 420
gatggttcaa tgaaaattca ggtcggcgca aatgatggtg aaaccatcac gattgatctg 480
aagaaaattg actctgatac gctgaatctg gctggtttta acgtgaatgg cgaagggtgaa 540
acagccaata ctgctgcaac acttaaagat atggttgggt taaaactcga taatagggg 600
gtcactacag ctggagttaa tagatatatt gctgacaaag ccgtcgcaag tagcacggat 660
attttgaatg cggtagctgg tgttgatggc agtaaagttt ccacggaggc agatgttgg 720
tttgggtcag ctgcccctgg tacgccagt gaataatact atcataaaga tactaacaca 780
tatacggctt ctgcttcagt tgatgcgact caactggcgg cattcctgaa tcctgaagcg 840
ggtggtacca ctgctgcaac agtaagtatt ggcaacggtg caacagctca agagcaaaaa 900
gtcattattg cttaaagatg ttctttaact gctgctgatg acgggtgccg tctctatctt 960
gatgatactg gtaacttaag taaaactaac gcaggcactg atactcaagc taaactgtct 1020
gacttaattg caaacaatgc taatgccaaa acagtcatta caacagataa aggtacattt 1080
actgctaata cgacaaagtt tgatggggta gatattttct ttgatgcttc aacgtttgct 1140
aacgccgtta aaaatgagac ttacactgca actggtgggt taactttacc tgcgacatat 1200
acagtcaata atggcactgc tgcactcagc tatttagtgc atggaaaagt gagcaaaact 1260

```

```

cctgccgagt attttgtctca agctgatggc actattacta gtggtgaaaa tgcggctacc 1320
agtaaaagcta tctatgtaag tgccaatggg aacttaacga ctaatacaac tagtgaatct 1380
gaagctacta ccaaccgcgt ggcagcattg gatgacgcta tcgcgtctat cgacaaattc 1440
cgttcttccc tgggtgctat ccagaaccgt ctggattccg cagtcaccaa cctgaacaac 1500
accactacca acctgtctga agcgcagtc cgtattcagg acgccgacta tgcgaccgaa 1560
gtgtccaaca tgtcgaaagc gcagatcatt cagcaggccg gtaactccgt gctggcaaaa 1620
gccaaaccagg taccgcagca ggttctgtct ctgctgcagg gttaa 1665

```

```

<210> 19
<211> 1842
<212> DNA
<213> Escherichia coli

```

```

<400> 19
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac cactgaaggc 240
gcgctgtccg aaattaacaa caacttacag cgtattcgtg aactgacggt tcaggcgacc 300
accggaacta actccacctc tgacctggac tccatccagg acgaaatcaa atcccgctct 360
gacgaaattg accgcgtatc tggtcagacc cagttcaacg gcgtgaacgt gctgtctaaa 420
gatggctcga tgaaaattca ggtcggcgcg aacgatggcg aaacgattac tattgatctg 480
aagaaaattg actctgatac gctgaatctg gctggtttta acgtaaacgg taaaggttct 540
gtagcgaata ccgctgcgac tacagataat ctgacattgg ctggttttac agcgggtact 600
aaagctgtcg atggcacctg aacttatagc aaaaatgtcc agtttgccgc cgcgactgca 660
agcaatgtac tggctgctgc taaagatggc gacgaaatta cgttcgctgg taataacggc 720
acaggtatag ctgcaactgg ggggacttat acttatcata aggactctaa ctcatacagc 780
tttagcgcaa cggctgcatac taaagattct ctggtgagca cactggcacc aaacgctggc 840
gatacattta ccgctaaagt gactattggg tctaaatcgc aagaagttaa cgtagcaaa 900
gatggtagca ttacatccag cgtgggtaag gcgctgtatt tagatgagaa gggcaacctg 960
acccaaacag gtagtggcac aaccaagct gcaacctggg ataacctgat ggccaatata 1020
gatactacag gcaaatgac ctatggtaac tctgcggcag cagctgttgg gacagtaatc 1080
gaagcaaaaag gaatgaccat cacttctgct ggtggtaatg ctcagggtgtt aaaagacgcg 1140
gcttataatg ccgcatatgc gacctcaatt actactggta ctccgggtga tgcgggagcc 1200
gcgggagccg ctgcaactgc gggtaatgcc gcggtgggag cgtggggcgc aacggcagtt 1260
gataatacca cggcagatgt tgccgatata tctatctcag cttcgcaaat ggcgagcatc 1320
cttcaggata aagatttcac cttaagtgat ggtagtgata cttacaacgt gaccagcaat 1380
gctgtcacta tcaatggcaa agcagcaaac attgatgaca gcggcgcaat cacagaccaa 1440
accagtaaaag ttgtcaatta ttctgctcat actaacggta gcgtgactaa cgatacaggc 1500
tccactattt atgcgacaga agatggtagc ctgaccaccg atgcagcaac caaagccgaa 1560
accaccgcgc atccctgtaa agctctggac gaagccatca gctccatcga caaatccgc 1620
tctccctcgt gtgcggtgca aaacgctctg gattccgcgg tcaccaacct gaacaacacc 1680
accaccaacc tgtctgaagc gcagtcccggt attcaggacg ccgactatgc gaccgaagtg 1740
tccaacatgt cgaaagcgca gattatccag caggccggta actccgtgct ggcaaaagct 1800
aaccaggtac cacagcagggt tctgtctctg ctgcagggtt aa 1842

```

```

<210> 20
<211> 1731
<212> DNA
<213> Escherichia coli

```

```

<400> 20
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg cggcccgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtccg aaattaacaa caacttacag cgtgtgcgtg agctgactgt tcaggcgacc 300
accggtacca actcccagtc tgatctggac tctatccagg acgaaatcaa atcccgctct 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt gctggcaaaa 420
gacggttcca tgaaaattca ggttggcgcg aatgatggcc agaccatcac tatcgacctg 480
aagaagattg actcttctac gttgaaactg actggtttta acgtgaatgg ttctggttct 540
gtggcgaata ctgcggcgac taaagcggat ttggctgctg ctgcaattgg taccctggg 600
gcagcagatt ctacaggtgc cattgcttac acagtaagtg ctgggctgac taaaactaca 660
gcgcgagatg tactgtctag cctcgctgat ggtacgacta ttacagccac aggcgtgaaa 720
aatggctttg ctgcaggagc cacttccaat gcctataaac ttaacaaaga taataatata 780

```

tttactttatg	acacgactgc	tacgacagct	gagctgcagt	cttacctgac	tccgaaagcg	840
ggcgacactg	caacattcag	tgttgaaatt	ggtgggtacta	cacaagacgt	cggtgctgtcc	900
agtgatggca	aactcactgc	taaggatggc	tctaagcttt	acattgatac	aactgggtaat	960
ttaactcaga	atgggtggtaa	taacgggtgt	ggaacactcg	cggaagcgac	tctgagtggg	1020
ttagctctga	acaaaaatgg	tttaacggct	gttaaatcca	caattactac	agctgataac	1080
acttcgattg	tactgaatgg	ttcaagcgat	ggtactggta	atgctggtac	tgaagggtacg	1140
attgctgtta	cagggcgtgt	aattagttca	gctgctctgc	aatctgcaag	caaaacgact	1200
ggtttcactg	ttggtacagt	agacacagct	ggttatatct	ctgtaggtag	tgatgggagt	1260
gttcaggcat	atgatgctgc	gacttcttgc	aacaaagctt	cttacaccaa	cactgacggg	1320
acactgacta	ctgataaac	cactaaactg	tatctgcaga	aagatggctc	tgtaaccaac	1380
ggttcaggta	aagcgggtcta	tgtagaagcg	gatgggtgatt	tcactaccga	cgctgcaacc	1440
aaagccgcaa	ccaccaccga	tccgctggcc	gctctggatg	acgcaatcag	ccagatcgac	1500
aagttccgtt	catccttggg	tgctatccag	aaccgtctgg	attctgcagt	caccaacctg	1560
aacaacacca	ccaccaacct	gtctgaagcg	cagtcgccgt	ttcaggacgc	cgactatgcg	1620
accgaagtgt	ccaatatgtc	gaaagcgag	atcatccagc	aggccggtaa	ctccgtgctg	1680
gcaaaagcca	accagggtacc	gcagcagggt	ctgtctctgc	tgcagggtta	a	1731

<210> 21

<211> 1380

<212> DNA

<213> Escherichia coli

<400> 21

aacaaatctc	agtcttctct	gagctccgcc	attgaacgtc	tctcttctgg	cctgcgtatt	60
aacagtgtca	aagatgacgc	agcagggtcag	gcgattgcta	accgttttac	agcaaatatt	120
aaaggtctga	ctcaggcttc	ccgtaacgcg	aatgatggta	tttctgttgc	gcagaccact	180
gaaggtgctc	tgaatgaaat	taacaacaac	ctgcagcgta	ttcgtgaact	ttctgttcag	240
gcaactaacg	gtactaactc	tgacagcgat	ctttcttcta	tccaggctga	aattactcaa	300
cgtctggaag	aaattgaccg	tgtatctgag	caaactcagt	ttaacggcgt	gaaagtcttt	360
gctgaaaata	atgaaatgaa	aattcagggt	ggtgctaattg	atgggtgaaac	catcactatc	420
aatctggcaa	aaattgatgc	gaaaactctc	ggcctggacg	gttttaatat	cgatggcgcg	480
cagaaagcaa	ccggcagtg	cctgatttct	aaatttaaag	cgacaggtag	tgataattat	540
caaattaacg	gtactgataa	ctatactgtt	aatgtagata	gtggcgtagt	acaggataaa	600
gatggcaaac	aagtttatgt	gagtactgcg	gatggttcac	ttacgaccag	cagtgtact	660
caattcaaga	ttgatgcaac	taagcttgca	gtggctgcta	aagatttagc	tcaagggaat	720
aagattgtct	acgaagggtat	cgaatttaca	aataccggca	ctgtcgctat	agatgccaaa	780
ggtaaatggta	aattaaccgc	caatgttgat	ggtaaggctg	ttgaattcac	tatttcgggg	840
agtactgata	catcaggtag	tagtgcaacc	gttgccctta	cgacagccct	atacaaaaaa	900
agtgacgggc	aattgactgc	aacaaaagtt	gaaaataaag	cagcgacact	atctgatctt	960
gatctgaacg	ctgccaagaa	aacagggaagc	acgttagttg	ttaacgggtg	aacttacgat	1020
gttagtgtag	atggtaaaa	gataacggag	actgcttctg	gtaacaataa	agtcattgtat	1080
ctgagcaaat	cagaagggtg	tagcccgatt	ctggtaaacg	aagatgcagc	aaaatcgttg	1140
caatctacca	ccaacccgct	cgaaaactatc	gacaaagcat	tggctaaagt	tgacaatctg	1200
cgttctgacc	tccgtgtagt	acaaaaccgt	ttcgactctg	ccatcaccaa	ccttggcaac	1260
accgtaaaaca	acctgtcttc	tgcccgtagc	cgatatcgaag	atgctgacta	cgcgacggaa	1320
gtgtctaaca	tgtctcgtgc	gcagatcctg	caacaagcgg	gtacctctgt	tctggcacag	1380

<210> 22

<211> 1767

<212> DNA

<213> Escherichia coli

<400> 22

atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagtctg	cgctgtcgag	ttctatcgag	cgctctgtct	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgcagcggg	tcaggcgatt	gctaaccgtt	ttacttctaa	cattaaaggc	180
ctgactcagg	cggcacgtaa	cgccaacgac	ggtatctctc	tggcgagac	caccgaagg	240
gcgctgtctg	aaatcaacaa	caacttacag	cgtgtactgt	aactgaccgt	tcaggcaacc	300

```

accggtacta actccgactc cgacctggct tctattcagg acgaaatcaa atccccgtctg 360
gatgaaattg accgcgtatc tggtcagact cagttcaacg gcgtgaacgt gctggcaaaa 420
gacggttcca tgaaaattca ggtaggtgct aacgacggcc agactatcac tattgacctg 480
aaaaaaaaatcg actctgatac tctgggcctg aatgggttta acgtgaatgg ttctgggacg 540
attaccaaca aagcagcaac tgtcagtgat gttactcgcg caggcgggtac attgggtgaat 600
ggtgcctatg atataaaaaac cactaacaca gcgctgacta caactgatgc cttcgcgaaa 660
ttgaatgatg gtgatgttgt tactatcaat aatggtaagg atactgccta taaatataat 720
gctgctacag gtgggttttac gacggatgtc tccatctccg gggatcctac cgctgctgac 780
gctactgcta ataaaactgc ccgtgatgca cttgcggcgt ctttacatgc tgagccgggt 840
aaaactgtta atgggttctg gactacgaat gatggtacgg taaaatttga taccgatgcc 900
gatggtaaga tttctattgg tgggttgct gcttatgtag atgcagcagg caacctgacc 960
actaacgcag caggatgatg gactcaagca acaactaccg atttgggttac tgctgctgca 1020
tctgctactg gtaaggggtg atccctgacc tttgggtgaca cgacgtataa aattgggtcag 1080
ggtacggctg ggggttgatcc tgatgacgct tcagatgatg tactgggcac catttcttac 1140
tctaaatcag taagcaagga tgttggttctt gctgatacta aagcaactgg taacacgaca 1200
acagttgatt tcaactccgg tatcatgact tcaaagggtta gtttcgatgc aggtacatca 1260
actgatacat tcaaagatgc agatgggtgct atcaccaaaa ctaaagaata caccacttct 1320
tatgctgtaa ataaagatac tgggtgaagt accgttgctg attatgctgc ggtagatagc 1380
gccgataagg ctggtgatga tactaaatat aaaccgacta tcggcgcgac agttaacctg 1440
aattctgcag gtaaattgac cactgatacc accagtgcag gcacagcaac caaagatcct 1500
ctggctgccc tggacgctgc tatcagctcc atcgacaaat tccgttcac cctgggtgct 1560
atccagaacc gtctggattc cgcagtcacc aacctgaaca acaccactac caacctgtcc 1620
gaagcgcagt cccgtattca ggacgccgac tatgcgaccg aagtgtccaa catgtcgaaa 1680
gcgcagatta tccagcaggc cggtaactcc gtgctggcaa aagccaacca ggtaccgcag 1740
caggttctgt ctctgctaca gggtaa 1767

```

```

<210> 23
<211> 1383
<212> DNA
<213> Escherichia coli

```

```

<400> 23
aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tttcttctg tctgcgtatt 60
aacagcgcta aagatgacgc tgcggggccag gcgattgcta accgcttcac ttctaaccatc 120
aaaggctctga ctcaggccgc acgtaacgcc aacgacggta tttctctggc gcagaccact 180
gaaggcgcgc tgcttgagat taacaacaac ttgcagcgtg tgcgtgagtt gactgtacag 240
gcgacgaccg ggactaactc tgattctgac ctgtcttcta tccaggatga aatcaaatcc 300
cgtttaagcg aaattgaccg tgatctcggc cagactcagt ttaacggcgt gaacgtactg 360
gctaagaatg acacctgtc tattcaggta ggtgcaaatg acggtcagac tatcaatatt 420
gacctgcagc aaatcgattc tcatacactg ggtctggatg gtttcagcgt taaaaataat 480
gatgcagtgga aaaccagtgc tgcctggaat actcttgggg ggggggcagg ttctgttgct 540
gtcgacttgc caacaaccag tttgactgct atcactggtc tcggtagcgg tgctatcagc 600
gaaattgcta aagacgataa tgggtgattac tacgcgcgat tcacagggac tacgggtaat 660
actgctgatg gttactatgc tgtcgatatac gacaaggcta ccggtgaggt cgctctgaaa 720
gatggtaacg tagatacacc gacaggtacg ccaacgacga caagcacata tgacttcaca 780
gacgtggctg aaaccgttcc ctttggcact gatgtgcaa cagccgggat cagcactggg 840
gcttctctcg ttaaaacttca ggatgagaaa ggcaatgata ctgctactta tgcaatcaaa 900
gcacaagatg gcagcctgta tgccgccaac gttgatgagg ctaccggtaa agtcactgtc 960
aaaaccgcca gctatactga tgctgacggc aaagcagtga ccgatgccgc tgtaaaactg 1020
ggtgggtgaca atggcacaaac cgaaaattgtt gtcgatgctg cgtcaggtaa aacttacgat 1080
gctgggtgcac tgcaaaacgt tgatctctcc agtgcaacca acacggtaac cgcaatcccg 1140
aacggtaaaa ccacgtctcc gctggctgcc cttgacgacg caatcagcca gatcgacaaa 1200
ttccgctcct cctcgggtgc ggtgcagaac cgtctggatt ccgcggtcac caacctgaac 1260
aacaccacta ccaacctgtc tgaagcgcag tcccgtattc aggacgctga ctatgcgacc 1320
gaagtatcca acatgtcgaa agcgcagatc atccagcagg caggtaactc cgtgctgtcc 1380
aaa 1383

```

```

<210> 24
<211> 1197
<212> DNA
<213> Escherichia coli

```

```

<400> 24
gcgctgtcga cttctatcga gcgcctctct tctggtctgc gcattaacag cgctaaagat 60
gacgctgcgg gccaaagcat tgctaaccgc ttcaacttca acatcaaagg tctgactcag 120
gccgcacgta acgccaacga cggtatctct ctggcgcgaga cactgaagg cgcactgtct 180
gaaatcaaca acaacttgca gcgtgttcgt gaactgaccg ttcaggccac taccggtact 240

```


aactctgatt	ctgacctgtc	ttcaatacag	gacgaaatca	aatcccgtct	cgatgaaatt	300
gaccgcgtat	ccggtcagac	tcagttcaac	ggcgttaatg	ttctttccaa	agatggttca	360
atgaaaattc	aggttggtgc	gaatgatggt	caaactatct	ccatcgatct	gaagaaaatt	420
gattcttcaa	ctttggggct	gaatggcttc	tcagtttcta	aaaactctct	taatgtcagc	480
aatgctatca	catctatccc	gcaagccgct	agcaatgaac	ctggtgatgt	taacttcggt	540
gatactgatg	agtctgcagc	aatcgagcc	aaattggggg	tttccgatac	gtcaagcctg	600
tcgctgcaca	acatccttga	taaagatggt	aaggcaacag	ctgattatgt	tgttcagtca	660
ggtaaagact	tctatgctgc	ttctgttaat	gccgcttcag	gtaaagtaac	cttaaacacc	720
attgatgtta	cttatgatga	ttatgcgaac	ggtggtgacg	atgccaaagca	aacagggtcag	780
ctgatcaaag	tttcagcaga	taaagacggc	gcagctcaag	gttttgtcac	acttcaaggc	840
aaaaactatt	ctgctgggtga	tgccgcagac	attcttaaga	atggagcaac	agctcttaag	900
ttaactgatc	tgaatttaag	tgatgttact	gataactaatg	gtaaggtaac	cacaactgcg	960
actgagcaat	ttgaaggtgc	ttcaactgag	gatccgctgg	cgcttctgga	taaagctatt	1020
gcacagtcg	acaaattccg	gtcttctcta	ggtgccgtgc	agaaccgtct	cgattccgct	1080
atcaccaacc	tgaacaacac	caccaccaac	ctgtctgaag	cgagtcctcg	tattcaggac	1140
gccgactatg	cgaccgaagt	gtccaacatg	tcgaaagcgc	agatcatcca	gcaggca	1197

<210> 25

<211> 1674

<212> DNA

<213> Escherichia coli

<400> 25

atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagtctg	cgctgtcgag	ttctatcgag	cgtctgtctt	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgccgcagg	tcaggcgatt	gctaaccggt	ttacttctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaacgac	ggtatttctg	ttgcacagac	cactgaaggc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgtattcgtg	aactgacggt	tcaggccact	300
acagggacta	actccgattc	tgacctggac	tccatccagg	acgaaatcaa	atctcgtctg	360
gacgaaattg	accgcgtatc	tggtcagacc	cagttcaacg	gcgtgaacgt	gctgtctaaa	420
gatggctcga	tgaaaattca	ggtcggcgcg	aacgatggcg	aaacgattac	tattgatctg	480
aagaaaattg	actctgatac	gttaaactctg	gctggtttta	acgtgaatgg	tgctggctct	540
gttgataatg	ccaaggcgac	tgccaaagat	cttactgatg	ctggttttac	ggcaagcgca	600
gctgatgcta	atggcaaaat	cacttatacc	aaagacaccg	ttactaaatt	cgacaaagcg	660
acagcggtcg	atgtattggg	caaagcggct	gctggcgata	gcattaccta	tgccggcact	720
gatactggct	taggagtcgc	tgctgatgcc	tcgacttaca	cctacaatgc	agccaataag	780
tcttacactt	ttgatgctac	tggtgttgcc	aaggcggatg	ctggaacggc	actgaaaggg	840
tacttaggcg	catctaacac	cggtaaaatt	aatatcggtg	gtaccgagca	agaagttaac	900
attgccaaag	atggctccat	caccgatacc	aatggcgatg	cgctgtatct	cgatagtacc	960
ggcaacttaa	ccaaaaatac	cgcgaatttg	ggggctgctg	ataaagcaac	tgtagataaa	1020
ctgttttgctg	gtgctcagga	tgcaacgata	accttcgata	gcggcatgac	agctaaattc	1080
gatcaaaactg	ctggtaaccgt	tgatttcaaa	ggcgcgtcta	tttctgctga	tgcaatggca	1140
tcaaccttaa	ataatggttc	ctatacagcc	aacgtagggtg	gtaaggctta	tgccgtaacc	1200
gctggcgag	ttcagacagg	tgccgcagat	gtgtataaag	ataccactgg	cgcactgacg	1260
actgaagatg	acgaaaccgt	taccgcgacc	tactacgggt	ttgctgatgg	taaagtttct	1320
gacggtggaag	gttctactgt	ctataaagct	gctgatgggt	ccatcactaa	agatgcgact	1380
accaagtctg	aagcaaccac	tgacctctg	aaagcccttg	acgacgcaat	cagccagatc	1440
gacaaattcc	gctcctccct	cggtgccgtt	caaaaccgtc	tggattccgc	cgtcaccaac	1500
ctgaacaaca	ccactaccaa	cctgtctgaa	gcgcagtcct	gtattcagga	cgccgactat	1560
gcgaccgaag	tgtccaacat	gtcgaaagcg	cagatcattc	agcaggccgg	taactccggt	1620
ctggcaaaaag	ccaaccagggt	accgcagcag	gttctgtctc	tgctgcaggg	ttaa	1674

<210> 26

<211> 1365

<212> DNA

<213> Escherichia coli

<400> 26

aacaaatctc	agtcttctct	tagctctgct	attgagcgtc	tctcttctgg	cctgcgtatt	60
aacagtgtcta	aagatgacgc	agcagggtcag	gcgattgcta	accgttttac	ggcaaatatt	120
aaaggtctga	ctcaggcttc	ccgtaacgcg	aatgatggta	tttctgttgc	gcagactact	180
gaagggtcgc	tgaatgaaat	taacaacaac	ctgcagcggtg	tacgtgaact	gactgttcag	240
gcaactaacg	gtactaactc	tgacagcgat	ctttcttcta	ttcaggcaga	aattactcaa	300
cgtctgggaag	aaattgaccg	tgtatctgag	caaactcagt	ttaacggcgt	gaaagtcctt	360
gccgaaaaat	atgaaatgaa	aattcagggt	ggtgctaattg	atggggaaac	catcactatc	420
aatctggcaa	aaattgatgc	gaaaactctc	ggcctggacg	gctttaatat	cgatggcgcg	480
cagaaagcaa	ctggcagtg	cctgatttct	aaatttaaag	cgacagggtac	tgataattat	540

```

caaattaacg gtactgataa ctatactggt aatgtagata gtggagcagt tcaaaatgag 600
gatgggtgacg caatttttgt tagcgctacc gatgggtctc tgactactaa gagtgatata 660
aaagtcggtg gtacaggtat tgatgcgact gggcttgcaa aagccgcagt ttcttttagct 720
aaagatgcct caattaaata ccaaggtatt actttcacca acaaaaggac tgatgcattt 780
gatggcagtg gtaacggcac tctaaccgct aatattgatg gcaaagatgt aacctttact 840
attgatgcga caggggaagga cgcaacatta aaaacgtctg atcctgttta caaaaatagt 900
gcaggtcagt tcaactacaac taaggttgaa aacaaagccg ctacagcatc ggatctggac 960
ttaaataacg ctaaaaaagt gggtagttct ttagttgtaa atggcgctga ttatgaagtt 1020
agcgctgatg gtaagacagt aactgggctt ggcaaaacta tgtatctgag caaatcagaa 1080
ggtaggtagcc cgattctggt aaaagaagat gcagcaaaat cgttgcaatc tactaccaac 1140
ccgctcgaaa ccacgcgaaa ggcattggct aaagttgaca atctgcgttc tgacctcggt 1200
gcagtacaaa accgtttcga ctctgctatc accaaccttg gcaacaccgt aaacaacctg 1260
tctctgccc gtagccgtat cgaagatgct gactacgcga ccgaagtgtc taacatgtct 1320
cgtgcgcaga tcttgcaaca agcgggtacc tctgttctgg cgcag 1365

```

<210> 27

<211> 1740

<212> DNA

<213> Escherichia coli

<400> 27

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagttct cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgat ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgatatccgtg aactgacggt tcaggcttct 300
accgggacta actccgattc ggatctggac tccattcagg acgaaatcaa atcccgctc 360
gacgaaattg accgcgtatc tggccagacc cagttcaacg gcgtgaacgt actggcgaaa 420
gacggttcaa tgaaaattca ggttgggtcg aatgacggcc agactatcac gattgatctg 480
aagaaaattg actctgatac gctggggctg agtgggttta atgtgaatgg tagcggggct 540
gtggctaata ctgcagcgac taaatctgat ttggcagcag ctcaactctt ggctccagg 600
actgctgatg ctaattggta agttacctat actggtggcg caggcctgaa aacatctaca 660
gctgcagatg taattgcgag tttggctaata aacgcaaaag ttaatgccac aattgcaaat 720
ggttttggat cgccaacagc tacagattat acatacaaca gcgctacagg cgattttaca 780
tatagtcaa ctattgcagc tggtaaaaat tctgggtgata gtaacagtgc tcagttacaa 840
tcttctctga caccaaaagc gggcgatact gctaacttaa acgttaaaat tggttctacg 900
tcaattgacg ttgtattggc tagcgacggg aaaattaccg cgaaagatgg ttcagaacta 960
tttattgacg tagatggtaa cctcactcaa aacaatgctg ggactgtcaa agcagccact 1020
cttgatgcac tgactaaaaa ctggcataca acaggcacac cgagtgcctt atctacggta 1080
attacaactg aagatgaaa aaaccttca ctggctggcg gtactgatgc tactacttct 1140
ggtgcaatca ctgtagcaaa tgcaagaatg agtgctgagt ctcttcaatc ggcaactaag 1200
tccacaggat tcacagttga tgttggagct actggtacca gcgcaggcga tattaagtt 1260
gatagttaa gtagatgaca acaacacaca ggtacagggt ttgaagacgc ttacaccaaa 1320
gctgatggtt cactgactac cgataataca accaatctgt ttttgcaaaa agacggaact 1380
gtgaccaatg gttcaggtaa agcagttctt gtttcagcgg atggttaatt tactactgac 1440
gctgaaaact aagctgcaac caccgcgat ccaactgaaag ctctggacga agcgatcacg 1500
tccatcgaca aattccgttc ttcctcggt gcggtgcaaa accgtctgga ttccgcagtc 1560
accaacctga acaacaccac tactaacctg tctgaagcgc agtcccgtat tcaggacgct 1620
gactatgcga ccgaagtgtc caatatgtcg aaagcgcaga tcatccagca ggccggtaac 1680
tccgtgctgg caaaagctaa ccagggtacc cagcaggttc tgtctctgct gcagggttaa 1740

```

<210> 28

<211> 1233

<212> DNA

<213> Escherichia coli

<400> 28

```

aacaataaac agtctgcgct gtgcacttct atcgagcgcc tctcttctgg tctgcgcatt 60
aacagcgcta aagatgacgc tgcgggccag gcgattgcta accgcttcac ttctaacatc 120
aaaggtctga ctcaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180
gaaggcgcac tgtctgaaat caacaacaac ttgcagcggt ttcgtgagct gaccgttcag 240
gccactaccg gtactaactc tgattctgac ctgtcttcaa tccaggacga aatcaaatcc 300
cgtctcgatg aaattgaccg cgtatccggt cagactcagt tcaacggcgt gaacgtactg 360
gcaaaagata acaccatgaa gattcagggt ggtgcgaacg atggtcagac tatatccatc 420
gactgcaaaa aaatcgactc ttctactctt gggttgaaac gtttctccgt ttctaaaaat 480
gctctcgaaa ctagcgaagc gatcactcag ttgcccgaac gtgcgaatgc accaatcgct 540

```

gtgaagatgg	atgcgtctgt	tctgaccgat	cttaacatta	ctgatgcttc	cgctgtttcg	600
ctgcacaacg	taactaaagg	tgggtgcgca	acgtctactt	atgttggtca	gtatggcgat	660
aagagctatg	cagcatctgt	tgatgcggga	ggtagacgtaa	aactgaataa	agccgacgta	720
acataataacg	acgcagcaaa	tgggtgttacg	aatggccacc	agattggtag	tctggttcag	780
gttggtgctg	atgcaaaaca	tgatgcagtt	ggttttgtta	ccgtgcaggg	gaaaaactat	840
gttgctaatg	actcattagt	caatgcta	ggcgtgctg	gcgctgcagc	aactagagtt	900
acaattgatg	gtgatggtag	ccttgagct	aaccaggcta	aaattgaact	tagccaaaat	960
ggtgctactg	ctgcaacatc	agagttcgct	ggtgcttcaa	ccaacgatcc	actgactctg	1020
ctggacaaaag	ctatcgcatc	tgttgataaa	ttccgttctt	ctttgggggc	ggtacagaac	1080
cgctcgagct	ccgctgtaac	caacctgaac	aacaccacta	ccaacctgtc	tgaagcgag	1140
tcccgatttc	aggacgccga	ctatgcgacc	gaagtgtcca	acatgtcgaa	agcgcagatc	1200
atccagcagg	caggtaactc	cgtgctgtcc	aaa			1233

<210> 29

<211> 1713

<212> DNA

<213> Escherichia coli

<400> 29

atggcacaag	tcattaatac	caacagcctc	tgcgtgatca	ctcaaaaataa	tatcaacaag	60
aaccagtctg	cgctgtcgag	ttctatcgag	cgtctgtctt	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgcgcgagg	tcaggcgatt	gctaaccgtt	ttacttctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaacgac	ggtagttctg	ttgcacagac	cactgaaggc	240
gcgctgtccg	aatcaacaa	caacttacag	cgtattcggt	aactgacggt	tcaggcgagc	300
accggaacta	actccacctc	tgacctggac	tccattcagg	acgaaatcaa	atcccgtctt	360
gatgaaattg	accgcgtatc	cggccaaaacc	cagttcaacg	gcgtgaacgt	actgtcaaaa	420
gatggctcga	tgaaaattca	ggcggcgca	aatgatgggt	aaaccatcac	gattgatctg	480
aaaaagatcg	actcttctac	attgaagctg	accagcttca	atgttaacgg	taaaggcgct	540
gttgataatg	ctaaagccac	tgaagcagat	ctgaccgctg	cgggcttctc	ccaagggtgca	600
gtcgtcagtg	gcaacagcac	ctggactaaa	tctactgtta	ctacctttaa	tgcagcaaca	660
gctaccgacg	tgctggcaag	cgttagcggc	ggcagcacta	ttagcggtta	taccggtaca	720
aacaatggat	taggcgtagc	ggcttctact	gcataacct	acaacgcaac	cagcaagtct	780
tattcatttg	acgcaaccgc	acttaccaat	ggcgatggta	ctggggccac	cactaaagtt	840
gctgatgtgc	tgaaagccta	tgcagcaaac	ggtagataata	cggctcagat	ctccatcggc	900
ggaagcgctc	aggacgttaa	aattgccagc	gatggcacc	tgactgacgt	caatgggtgat	960
gctttatata	ttggttctga	cggaacacctg	actaaaaacc	aggccggcgg	tccagatgcg	1020
gcaacggttg	acggtatttt	caacgggtgcg	aatggtaatg	cagcagttga	tgcgaagatt	1080
acattcgcca	gcggcatgac	cgttgatttc	accaggcta	gcaaaaaagt	ggatattaag	1140
ggcgcaacgg	tatccgccga	agatatggac	actgcgttaa	ctgggcaggc	ttataccgta	1200
gctaaccggc	cacagtcttt	tgacgttgcc	gctgggtggg	cagtaaccgc	tactacaggt	1260
ggcgctaccg	taaaatttgg	tgctgatggg	gaactgacga	ctgcgaccaa	caagactgtc	1320
acagaaactt	atcacgaatt	tgctaaccggc	aatattctgg	atgatgacgg	cgcggctctg	1380
tacaaagcgg	ctgacggttc	tctgaccact	gaagctactg	gtaaatccga	agtgaccacg	1440
gatccgctga	aagcgtgga	cgatgctatc	gcattccgtg	acaaattccg	ctcctccctc	1500
ggtgcgggtg	agaaccgtct	ggattccgca	gtccaccaacc	tgaacaacac	cactaccaac	1560
ctgtctgaag	cgcagtcgcc	cattcaggac	tcggaatcaa	cgaaccgaag	gtccaatatg	1620
tcgaaagcgc	agatcatcca	gcaggccggg	aactccgtgc	tggcaaaaagc	caaccaggta	1680
ccgcagcagg	ttctgtctct	gctgcagggt	taa			1713

<210> 30

<211> 1668

<212> DNA

<213> Escherichia coli

<400> 30

atggcacaag	tcattaatac	caacagcctc	tgcgtgatca	ctcaaaaataa	tatcaacaag	60
aaccagtctg	cgctgtcgag	ttctatcgag	cgtctgtctt	ctggcttgcg	tattaacagc	120
gctaaggatg	acgcgcgagg	tcaggcgatt	gctaaccgtt	ttacttctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaacgac	ggtagttctg	ttgcgcagac	cactgaaggc	240
gcgctgtccg	aatcaacaa	caacttacag	cgtattcggt	aactgacggt	tcaggcttct	300
accgggacta	actccgattc	tgacctggac	tccattcagg	acgaaatcaa	atcccgtctg	360
gacgaaattg	accgcgtatc	tgccagagac	cagttcaacg	gcgtgaacgt	actggcgaaa	420
gacgggtcaa	tgaaaattca	ggttggtgcg	aatgacggcc	agactatcac	tattgatctg	480
aagaaaattg	actcagatac	gctggggctg	agtgggttta	atgtgaatgg	tggcggggct	540
ggtgctaata	ctgcagcgac	taaagatgat	ttggtcgctg	catcagtttc	agctgcggta	600
ggtaattgaat	acactgtctc	tgctggcctg	tcgaaatcaa	ctgctgctga	tggtattgct	660
agttctcacag	atgggtcgac	agtaactgcg	gctgggtgtaa	gcaatgggtt	tgctgcaggg	720

gcaactggag	atgcttataa	attcaatcaa	gcaaacaaca	cttttactta	caataaccacc	780
tcaacagcgg	cagaactcca	atcttacctc	acgcctaagg	cggggggatac	cgcaactttc	840
tccgttgaaa	ttggtggcac	caagcaggat	gttggttctg	ctagtgatgg	caaaatcaca	900
gcaaagacg	ggtctaaact	ttatattgac	accacagggg	atttaaccca	aaacgggtgga	960
ggtacttttag	aagaagctac	octcaatggc	ttagctttca	accactctgg	tccagccgct	1020
gctgtacaat	ctactattac	tactgcggt	ggaacttcaa	tagttctagc	aggttctggc	1080
gactttggaa	caacaaaaac	tgctggggct	attaatgtca	caggagcagt	gatcagtgtc	1140
gatgcacttc	tttccgccag	taaagcgact	gggtttactt	ctggcactta	taccgtaggt	1200
acagatggag	ttgttaaadc	tggtggcaat	gacgtttata	acaaagctga	cgggacggga	1260
ttactactg	acaataccac	aaaatattat	ttacaagatg	acgggtctgt	aactaatggt	1320
tctgggtaaa	ctgtgtatgc	tgatgcaaca	ggaaaactaa	ctactgacgc	tgaaaactaaa	1380
gccgaaacca	ccgccgatcc	cctgaaagct	ctggacgaag	cgatcagctc	catcgacaaa	1440
ttccgttctt	ccctcggtgc	ggtgcaaaac	cgtctggatt	ccgcggtcac	caacctgaac	1500
aacaccacta	ccaacctgtc	cgaagcgag	tcccgtattc	aggacgccga	ctatgcgacc	1560
gaagtgtcca	acatgtcgaa	agcgcagatc	atccagcagg	cggtaactc	cgtgctggca	1620
aaagctaacc	aggtaccgca	gcaggttctg	tctctgctgc	agggttaa		1668

<210> 31

<211> 1713

<212> DNA

<213> Escherichia coli

<400> 31

atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagttctg	cgctgtcgag	ttctatcgag	cgtctgtctt	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgcgcggg	tcaggcgatt	gctaaccggt	ttacttctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaacgac	ggatattccg	ttgcgcagac	caccgaaggc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgtatccgtg	aactgacggg	tcaggccact	300
accggtacta	actccgattc	tgacctggac	tccatccagg	acgaaatcaa	atctcgtctt	360
gatgaaattg	accgcgtatc	tggtcagacc	cagttcaatg	gcgtgaatgt	gttgtccaaa	420
gacggttcaa	tgaaaattca	ggtgggcgca	aatgatggg	aaaccatcac	gattgacctg	480
aaaaaaatcg	actcttctac	actgaagctg	accagcttca	acgtcaacgg	taaaggcgct	540
gttgataatg	caaaagccac	tgaagcagat	ctgaccgctg	cgggcttctc	ccaaagtgca	600
gttgtcagtg	gcaatagcac	ctggactaaa	tctactgtta	ctacctttaa	tgcaagcaaca	660
gctaccgatg	tgctggctag	cgttagtggt	ggcagcacta	ttagcgggta	tgctggcaca	720
aacaaatggg	taggcgtagc	ggcttctact	gcataacct	acaacgcaac	cagcaagtct	780
tattcatttg	acgcaaccgc	acttactaat	ggtgatggta	ctgcgggctc	aactaaagtt	840
gctgatgttc	tgaaagccta	tgcaagcaac	ggcgataaca	cggctcagat	ctccatcggt	900
ggtagcgctc	aggaagttaa	aattgccagc	gatggtaccc	tgacggatac	taattggcgat	960
gctttatata	ttggtgctga	cggtaacctg	acgaaaaacc	aggccggcg	cccagcccg	1020
gcaacggttg	acggtatttt	caacgggtgc	aatgggtcatg	atgcagttga	tgcaagatt	1080
accttcggca	gcggcatgac	cgttgacttc	accaggtta	gcaacaatgt	ggatattaag	1140
ggcgcgacgg	tatccgccga	agatatgaac	actgcgttaa	ccggtcaggc	ttataccgta	1200
gctaaccggc	cacagtctta	tgacgttgcc	gctgatggg	cagtaactgc	tactacaggt	1260
ggagcgacgg	taaatattgg	tgctgaggg	gaactgacga	ctgcggccaa	caagactgtc	1320
acagaaaact	atcacgaatt	tgctaaccgc	aatattctgg	atgatgacgg	cgcggctctg	1380
tataaagcgg	ctgacggctc	tctgaccact	gaagctacag	gtaaatctga	agcgaccacg	1440
gatccgctga	aagcgctgga	cgatgctatc	gcattccgtg	acaaattccg	ttcttccctg	1500
ggtgccgtgc	agaaccgtct	ggattccgca	gtcaccaacc	tgaacaacac	cactaccaac	1560
ctgtccgaag	cgcagtcctc	tattcaggac	gccgactatg	cgaccgaagt	gtccaacatg	1620
tcgaaagcgc	agattattca	gcaggcaggt	aactccgtgc	tggcaaaagc	taaccaggta	1680
ccgcagcagg	ttctgtctct	gctgcagggt	taa			1713

<210> 32

<211> 1188

<212> DNA

<213> Escherichia coli

<400> 32

aacaaaaacc	agtctgcgct	gtcgacttct	atcgagcgcc	tctcttctgg	tctgcgcatt	60
aacagcgcta	aagatgacgc	tcggggcccag	gcgattgcta	accgcttcac	ttctaacatc	120
aaaggtctga	ctcaggccgc	cgtaacgcc	aacgacggta	tctctctggc	gcagaccact	180
gaaggcgcac	tgtctgaaat	caacaacaac	ttgcagcgtg	tgctgaggtt	gactgttcag	240
gcgacgaccg	ggaactaactc	tgattctgac	ctgtcttcta	ttcaggacga	aatcaaatcc	300
cgtctggatg	aaattgaccg	tggttccggg	cagaccaggt	tcaacggcgt	gaacgtgctg	360
gctaaaaacg	gttctatggc	gattcaggtt	ggcgcaaatg	atgggcagac	catcaacatc	420
gacctgcaga	aaatcgactc	ttctactctg	ggcctgggcg	gcttctccgt	atctaacaat	480

```

gcactgaaac tgagcgattc tatcactcag gttggtgcga gtggttcact ggcagatgtg 540
aaactgagct ctgttgcttc ggctctgggt gtagacgcaa gcactctgac tctgcacaac 600
gtacagaccc cagctggcgc agcaacagct aactatgttg tctcttctgg ttctgacaac 660
tactcagtat ctgttggaaga tagctccggt acagttacgc tgaacaccac tgatataggt 720
tataccgata ccgctaattg cgttactacc ggttccatga ctggtaagta cgttaaagtt 780
ggagctgatg cattgggtgc tgctgtaggt tatgtcacgc tacagggaca aaacttcaaa 840
gctgatgctg gcgcgctggt taactccaag aatgctgctg gtagtcagaa tgttacttct 900
gcaattggcg atattgctaa taaagcgaat gctaacattt acactggaac ctcttctgca 960
gatccactgg ctctgctgga caaagctatc gcactctgtg ataaattccg ttcttctcta 1020
ggggcggtgc agaaccgtct gagctctgct gtaaccaacc tgaacaacac cactaccaac 1080
ctgtccgaag cgcagtcccg tattcaggac gccgactatg cgaccgaagt gtccaacatg 1140
tcgaaagcgc agatcatcca gcaggcgggt aactccgtgc tgtctaaa 1188

```

<210> 33
 <211> 1638
 <212> DNA
 <213> Escherichia coli

```

<400> 33
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtcctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgccgg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaatgac ggtatctctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggt tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
gacgaaaattg accgcgtatc cggctcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaaaattca ggttgggtgc aacgacggcc agactatcac tattgatctg 480
aagaaaattg actctgatac gctggggctg agtgggttta acgtaaatgg tagcgcagat 540
aaggcaagtg tcgcggcgac agctgacgga atgggttaaag acggatatat caaagggtta 600
acttcatctg acggcagcac tgcatatact aaaactacag caaatactgc agcaaaagga 660
tctgatattc ttgcggcgct taagactggc gataaaatta ccgcaacagg tgcaaatagc 720
cttgctgata atgcgacatc gacaacttat acttataatg caaccagcaa taccttctcc 780
tatacggctg acggtgtaaa ccaaacgaat gctgcagcaa atctcatacc tgcagcaggg 840
aaaacgacag ctgcatcagt tactattggt gggacagcac agaattgtaa tattgatgat 900
tcgggcaata ttacttcaag tgatggcgat caactttatc tggattcaac aggtaacctg 960
actaaaaacc aggccggcaa cccgaaaaaa gcaaccgttt ctgggcttct cggaaatacg 1020
gatgcgaaag gtactgctgt taaaacaacc atcaagacag aggctggtgt aacagttaca 1080
gctgaaggta atacaggta cgtaaaaatt gaagggtgta ctggttcagc atctgcattt 1140
acgggcattg catattccgc caacaccggt ggggaatact atgctgttgc cgcaataat 1200
actacaaatg ttctctggc gggggatgac ttaacccagg atgctcaaac tgtttcaacc 1260
tactactcgc aagccgatgg caccggtcac aatagcgcag gcaaagaaat ctataaagac 1320
gctgatgggtg tctacagcac agagaataaa acatcgaaga cgtccgatcc attggctgcg 1380
cttgacgacg caatcagctc catcgacaaa ttccgttcat ccttgggtgc tatccagaac 1440
cgtctggatt ccgcggtcac caacctgaac aacaccacta ccaacctgtc cgaagcgcag 1500
tcccgtattc agacgcgca ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1560
atccagcagg ccgtaactc cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1620
tctctgctgc agggctaa 1638

```

<210> 34
 <211> 2145
 <212> DNA
 <213> Escherichia coli

```

<400> 34
aacaatctc agtcttctct gagctccgcc attgaacgtc tctcttctgg cctgcgtatt 60
aacagtgtc aagatgacgc agcagggtcag gcgattgcta accgttttac agcaaatatt 120
aaaggtctga ctcaggcttc ccgtaacgcg aatgatggta tttctgttgc gcagaccact 180
gaaggtgcgc tgaatgaaat taacaacaac ctgcagcgtg tacgtgaact gactgttcag 240
gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggctga aattactcaa 300
cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360
gctgaaaata atgaaatgaa aattcagggt ggtgctaatt atgggtgaaac catcactatc 420
aatctggcaa aaattgatgc gaaaactctc ggctggacg gttttaatat cgatggcgcg 480
cagaaagcaa ctggcagtg cctgatttct aaatttaaag cgacagggtac tgataactat 540
gatgttggcg gtgatgctta tactgttaac gtagatagcg gagctgggta atgactccaa 600
cttattgata tggttttatg ttcagataat gcccgatgac tttgtcatgc agctccaccg 660
atthttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc agattcagggt 720

```

tatgccgctc	aattoegctgc	gtatatcgct	tgctgattac	gtgcagcttt	cccttcaggc	780
gggattcata	cagcggccag	ccatccgtca	tccatatcac	cacgtcaaag	ggtgacagca	840
ggctcataag	acgccccagc	gtcgccatag	tgcgttcacc	gaatacgtgc	gcaacaaccg	900
tcttcgggag	cctgtcatac	gcgtaaaaca	gccagcgctg	gcgcgattta	gccccgacat	960
agtcccactg	ttcgtccatt	tccgcgcaga	cgatgacgtc	actgcccggc	tgtagcgcg	1020
aggttaccca	ctgcggcctg	agttttttaa	gtgacgtaaa	atcgtgttga	ggccaacgcc	1080
cataatgcgg	gcagttgccc	ggcatccaac	gccattcatg	gccatatcaa	tgattttctg	1140
gtgctgaccg	ggttgagaag	cgggtgtaagt	gaactgcagt	tgccatgttt	tacggcagtg	1200
agagcagaga	tagcgctgat	gtccggcggt	gcttttgccg	ttacgcacca	ccccgtcagt	1260
agctgaacag	gagggacagc	tgatagaaac	agaagccact	ggagcacctc	aaaaacacca	1320
tcatacacta	aatcagtaag	ttggcagcat	taccgcggag	ctgttaaaga	tactacaggg	1380
aatgatattt	ttgttagtgc	agcagatggt	tactgacaa	ctaaatctga	cacaaacata	1440
gctggtacag	ggattgatgc	tacagcactc	gcagcagcgg	ctaagaataa	agcacagaat	1500
gataaattca	cgtttaatgg	agttgaattc	acaacaacaa	ctgcagcggg	tggcaatggg	1560
aatggtgtat	attctgcaga	aattgatggt	aagtcagtga	catttactgt	gacagatgct	1620
gacaaaaaag	cttctttgat	tacgagtga	acagtttaca	aaaatagcgc	tggcctttat	1680
acgacaacca	aagttgataa	caaggctgcc	acactttccg	atcttgatct	caatgcagct	1740
aagaaaacag	gaagcacgtt	agttgttaac	ggtgcaactt	acgatgttag	tgcatgtggt	1800
aaaacgataa	cggagactgc	ttctggtaac	aataaagtca	tgtatctgag	caaatcagaa	1860
ggtggtagcc	cgattctggt	aaacgaagat	gcagcaaaat	cgttgcaatc	taccaccaac	1920
ccgctcgaaa	ctatcgacaa	agcattggct	aaagttgaca	atctgcgttc	tgacctcggt	1980
gcagtacaaa	accgtttcga	ctctgctatc	accaaccttg	gcaacaccgt	aaacaacctg	2040
tcttctgccc	gtagccgtat	cgaagatgct	gactacgcga	ccgaagtgtc	taacatgtct	2100
cgtgcgcaga	tcctgcaaca	agcgggtacc	tctgttctgg	cgcag		2145

<210> 35
 <211> 1587
 <212> DNA
 <213> Escherichia coli

aacaagaacc	agtctgcgct	gtcgagttct	atcgagcgctc	tgtcttctgg	cttgcgtatt	60
aacagcgcta	aggatgacgc	cgcaggtcag	gcgattgcta	accgtttttac	ttctaacatt	120
aaaggcccta	ctcaggctgc	acgtaacgcc	aacgacggta	tttctgttgc	gcagaccacc	180
gaaggcgctg	tgctcgaaat	caacaacaac	ttacagcgctg	tgctgtaact	gaccgttcag	240
gcaaccaccg	gtaccaactc	ccagctcgac	ctggactcta	tccaggacga	aattaaatcc	300
cgtctggacg	aaattgaccg	cgtatccggt	cagaccagct	tcaacggcgt	gaacgtactg	360
gcaaaagacg	gttccatgaa	aattcaggtt	ggcgcgaaacg	atggccagac	catcactatc	420
gacctgaaga	agattgactc	ttctacgctg	aaactgactg	gttttaacgt	gaatggcaaa	480
gcagcggttg	ataatgctaa	agcgacggat	gcaaactctga	ctaccgccgg	ttttacacaa	540
ggcggttggtg	attcaaatgg	taatagtact	tggaactaaat	caactacgac	taatttcgat	600
gcggcaactg	cagtaaacgt	actagcagca	gttaaagatg	gcagcacaaat	caattacacc	660
ggtactggta	atgggtttag	gattgctgca	acaagtgtct	atacatatca	cgatagcact	720
aaatcctata	cctttgatct	tacgggggct	gcagtagctg	gtgccgcgctc	cagcctgcaa	780
ggtacttttg	gtacagatac	gaatactgca	aaaatcacca	tcgatggttc	tgctcaagaa	840
gtaaacatcg	ctaaagatgg	gaaaattact	gatactgatg	gtaaagcttt	atatatcgat	900
tccactggta	atttgactaa	gaacggctct	gatactttta	ctcaggcaac	attgaatgat	960
gtccttactg	gtgctaattc	agttgatgat	acaaggattg	acttcgatag	cgcatgtct	1020
gtcacccttg	ataaagtga	cagcactgta	gatatcactg	gcgcactctat	ttcagccgct	1080
gcaatgacta	atgagttgac	aggtaaggcc	tataccgtag	taaatgggtg	agaatcttac	1140
gctgtagcta	ctaataacac	agtaaaaacg	actgctgatg	ctaaaaatgt	ttatgttgat	1200
gctagtggta	aattaactac	tgatgacaaa	gccactgtta	cagaaactta	tcatgaattt	1260
gcgaatggca	atatctatga	tgataaaggc	gctgctgttt	atgcggcggc	ggatgggtct	1320
ctgactacag	aaactacaag	taaatacaga	gctacagcta	acccgctggc	cgctctggac	1380
gacgcaatca	gccagatcga	caaattccgt	tcattccctg	gtgctatcca	gaacgctctg	1440
gattccgcag	tcaccaacct	gaacaacacc	ataccaatc	tgtctgaagc	gcagtcctcg	1500
attcaggacg	ccgactatgc	gaccgaagtg	tccaatatgt	cgaaagcgca	gatcatccag	1560
caggcaggca	actccgtgct	ggcaaaa				1587

<210> 36
 <211> 1245
 <212> DNA
 <213> Escherichia coli

aacaaaaaac	agtctgcgct	gtcgacttct	atcgagcgcc	tctcttctgg	tctgcgcatt	60
aacagcgcta	aagatgacgc	tgccggccag	gcgattgcta	accgcttcac	ttctaacatc	120

```

aaaggtctga ctcaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180
gaaggcgcac tgtctgaaat caacaacaac ttgcagcgtg ttcgtgaact gaccgttcag 240
gccactaccg gtactaactc tgattctgac ctgtcttcaa tccaggacga aatcaaattcc 300
cgtctcgatg aaattgaccg cgtatccggt cagactcagt tcaacggcgt gaacgtactg 360
gcaaaagatg gctcgatgaa aattcaggtc ggtgcaaatg atggtcagac aatcagcatt 420
gatttgcaga agattgattc ttctacttta gggttaaatg gttttctgtt ttccaaaaat 480
gcagtatctg ttggtgatgc tattactcaa ttgcctggcg agacggcagc cgatgcacca 540
gtaaccatca agtttgatga ttcagtaaaa actgatttaa aactgaccga tgcttcaggg 600
ttaagtctgc ataacctcaa agatgaaaat ggtaatttaa ctaaccagta tgttgtagag 660
aatggcgcaa aatcttacgc tgctacagtc gctgccaatg gtaatgttac gctgaacaaa 720
gcaaatgtaa cctacagcga tgctgcaaac ggtattgata ccgcaacgca gtcaggccag 780
ttagttcagg ttggtgcaga ttctaccggt acgcccacaa cttcgtgtc tgccaagggt 840
aaaagctttg gcattgatga cgcgccttg aagaataaca ctggtgatgc taccgctact 900
caaccgggaa catctgggac aacagttgtc gcagcgtcaa ttcattctgag tacgggcaaa 960
aactctgtag acgctgatgt aacggcttcc actgaattca cagggtgttc aaccaacgat 1020
ccactgactc tgctggacaa agctatcgca tctgttgata aattccgttc ttctttgggg 1080
gcggtacaga accgtctgag ctccgctgta accaactga acaacaccac caccaacctg 1140
tctgaagcgc agtcccgtat tcaggacgcc gactatgcga ccgaagtgtc caacatgtcg 1200
aaagcgcaga ttatccagca ggcaggtaac tccgtgctgt ccaaa 1245

```

```

<210> 37
<211> 1185
<212> DNA
<213> Escherichia coli

```

```

<400> 37
aacaaaaacc agtctgcgct gtcgaattct atcgagcgcc tctcttctgg tctgcgcatt 60
aacagcgcta aagatgacgc tgcgggccag gcgattgcta accgcttcac ttctaacatc 120
aaaggtctga ctcaggctgc acgtaacgcc aatgacggta tttctctagc acagacagcg 180
gaaggcgcgc tgtcagagat taacaacaac ttgcagcgtg tgcgtgagtt gaccgtgcag 240
gcaaccactg gtaccaactc tgattccgat ctctcttcta ttcaggatga aattaaatct 300
cgtctggatg aaattgaccg cgtctctggt cagacccagt ttaacggcgt gaacgtactg 360
gctaaaaacg gttctatggc aattcagggt ggcggaacg atggccagac tatctctatc 420
gacctgcaga aaatagactc ttctactctg ggtctgagcg gcttctctgt ttctcagaac 480
tccctgaaac tgagcgattc tatcactacg atcggaata ctactgtgc atcgaagaac 540
gtggacctga ggcagtagtc aactaaactg gcggtgaatg caagcaccct gagcctgcac 600
gaagttcagg actctgctgg tgacggtact ggtaccttcg ttgtttcttc tggcagcgac 660
aactatgctg tgtctgtaga cgcggcctct ggtgcagtta acctgaacac cactgacgtc 720
acctatgatg acgctactaa tgggtgtact ggcgcgactc agaacgggtc gctgatcaaa 780
gtaacttctg acgccaacgg tgcagctgtt ggttacgtaa ccattcaggg taaaaactat 840
caggctggty gcaccggtgt tgacgttctg gcgaacagcg gtgttgacg ctaactactga gctgaaaggt 960
gctgttgata ccggtactct gcaactgagc ggtactggty caactactga gctgaaaggt 960
actgcaactc agaaccact ggcactattg gacaaagcta tcgcttctgt tgataaattc 1020
cgttcttctc tgggtgcggt acagaatcgt ctgagctctg ctgtaaccaa cctgaataac 1080
accaccacta acctgtctga agcgcagctc cgtattcagg atgccgacta tgcgaccgaa 1140
gtgtcaaata tgtctaaagc gcagatcggt cagcaggccg gtaac 1185

```

```

<210> 38
<211> 1383
<212> DNA
<213> Escherichia coli

```

```

<400> 38
aacaaatctc agtcttctct tagctctgct attgagcgtc tgtcttctgg tctgcgtatt 60
aacagcgcaa aagacgatgc agcaggtcag gcgattgcta accgttttac ggcaaatatt 120
aaaggtctga cccaggcttc ccgtaacgca aatgatggta tttctgttgc gcagaccact 180
gaagggtgcy tgaatgaaat taacaacaac ctgcagcgtg ttcgtgaact ttctgttcag 240
gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggctga aattactcaa 300
cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360
gctgaaaata atgaaatgaa aattcagggt ggtgctaata atggtgaaac catcactatc 420
aatctggcaa aaattgatgc gaaaactctc ggcctggacg gttttaatat cgatggcgcg 480
cagaaagcaa caggcagtga cctgatttct aaatttaaag cgacagggtac tgataattat 540
gatgttggcg gtaaaactta taccgtgaat gtggagagcg gcgcgggtta gaatgatgct 600
aataaagatg tttttgtaag cgcagctgat ggatcgctga cgaccagtag tgatactaaa 660
gtatccggtg aaagtattga tgcaacagaa ctacggaac ttgcaataaa attagctgac 720
aaaggctcca ttgaatacaa gggcattaca ttactaaca acactggcgc agagcttgat 780
gctaattggta aagggtgtttt gaccgcaaat attgatggty aagatgttca atttactatt 840

```

gacagtaatg	caccacg	tgccggcgca	acaataacta	cagacacagc	tgtttacaaa	900
aacagtgcgg	gccagttcac	cactacaaaa	gtggaaaata	aagccgcaac	actctctgat	960
ctggatctta	atgcagccaa	gaaaacaggt	agcactttag	ttgtaaatgg	cgccacctac	1020
aatgtcagcg	cagatggtaa	aacggtaact	gatactactc	ctggtgcccc	taaaagtgatg	1080
tatctgagca	aatcagaagg	tggtagcccg	attctggtaa	acgaagatgc	agcaaaatcg	1140
ttgcaatcta	ccaccaaccc	gctcgaaact	atcgacaagg	cattggctaa	agttgacaat	1200
ctgcgttctg	acctcggtgc	agtacaaaac	cgtttcgact	ctgccatcac	caaccttggc	1260
aacaccgtaa	acaacctgtc	ttctgcccgt	agccgtatcg	aagatgctga	ctacgcgacc	1320
gaagtgtcta	acatgtctcg	tgcgagatc	ctgcaacaag	cgggtacctc	tgttctggcg	1380
cag						1383

<210> 39
 <211> 1680
 <212> DNA
 <213> Escherichia coli

atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagtctg	cgctgtcgag	ttctatcgag	cgtctgtctt	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgccgcagg	tcaggcgatt	gctaaccgtt	tcacctctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaacgac	ggatatttctg	ttgcacagac	caccgaaggc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgtatccgtg	aactgacggg	tcaggcttct	300
accggggacta	actctgattc	ggatctggac	tccattcagg	acgaaatcaa	atcccgtctg	360
gacgaaattg	accgcgtatc	cggccagacc	cagttcaacg	gcgtgaacgt	gctggcgaaa	420
gacgggtcaa	tgaaaattca	ggttggtgcg	aatgacggcc	agactatcac	tattgatctg	480
aagaaaattg	actctgatac	tctggggttg	agtggattta	atgtgaatgg	caaaggggct	540
gtggctaacg	caaaagcgac	cgaagcagat	ttaacggggg	ctgggtttctc	tcaaggagcg	600
gtggatacaa	acggaaatag	tacttggaac	aaatcaacca	ccaccaatta	ctcagctgca	660
acaactgctg	acttgttatc	gaccattaag	gatggctcta	ctgttacata	tgcagggaca	720
gacaccggat	taggggtcgc	agcagcagga	aattatactt	atgatgcgaa	cagtaaatct	780
tattccttca	atgccaatgg	tctgacgggc	gcaaataccg	caactgcact	caaagggttac	840
ttggggacag	gtgctaacac	cgctaaaatt	tctatcgggtg	gtacagagca	ggaagtgaat	900
attgccaaag	atggcactat	tacagatacg	aatgggtgatg	cgctctatct	ggatattacc	960
ggcaacctga	ctaagaacta	tgccgggttca	ccacctgcag	caacgcgtgga	taacgtatta	1020
gcttccgcaa	ctgtaaatgc	cactatcaag	tttgatagcg	gtatgacggg	tgattacact	1080
gcaggtactg	gcgcgaatat	tacagggtgca	tccatttctg	cagatgacat	ggccgcgaaa	1140
ctgagcggaa	aggcgtacac	tgttgccaat	ggtgctgagt	cttatgacgt	tgctgcagtt	1200
acgggggctg	taacaactac	agcaggtaat	tcacctgtgt	atgccgatgc	agacggtaaa	1260
ttaacgacga	gtgccagtaa	tacggttact	cagacttatc	acgagtttgc	taatggtaac	1320
atztatgatg	acaaaggctc	gtcactgtat	aaagctgcag	atggctctct	gactttcgaa	1380
gctaaaggga	aattctgaagc	aaccgcgat	cccctgaaaag	ctctggacga	agccatcagc	1440
tccatcgaca	aattccgctc	ctccctcggt	gccgttcaaa	accgtctgga	ttctgcggtg	1500
accaacctga	acaacaccac	taccaacctg	tctgaagcgc	agtcccgtat	tcaggacgcc	1560
gactatgcga	ccgaagtgtc	caatatgtcg	aaagcgcaga	tcattccagca	ggccggtaac	1620
tccgtgttg	caaaagctaa	ccagggtaccg	cagcaggttc	tgtctctgct	gcagggttaa	1680

<210> 40
 <211> 1146
 <212> DNA
 <213> Escherichia coli

gcgctgtcga	cttctatcga	gcgcctctct	tctggtttgc	gcattaacag	cgctaaagat	60
gacgtcgcgg	gccaggcgat	tgctaaccgc	ttcacttcta	acatcaaagg	tctgactcag	120
gccgcacgta	acgccaacga	cggatatctc	ctggcgcaga	ccactgaagg	cgcactgtct	180
gaaatcaaca	acaacttgca	gcgtgttcgt	gaactgaccg	ttcaggccac	taccggtact	240
aactctgatt	ctgacctgtc	ttcaatccag	gacgaaatca	aatcccgtct	ggctgaaatc	300
gatcgtgtct	ctggtcagac	ccagttcaac	ggcgtgaacg	tgctggctaa	aaacggttct	360
ctgaatattc	aggttggcgc	gaatgatggg	cagaccatct	ctatcgattt	gcagaaaata	420
gactcttctg	cccttggttt	aagtggtttt	agtgttgccg	gtggggcgct	aaaattaaagc	480
gatacagtga	cgcaggtcgg	cgatggttca	gccgcgccag	ttaaagtggg	tctggatgca	540
gcagcaacag	atattggtac	tgctttgggg	caaaagggtta	atgcaagttc	tttaacggtg	600
cacaatatct	tagacaaaga	tggtgcggca	actgagaact	atgttggttag	ctatggtagt	660
gataattacg	ctgcatctgt	tgacagatgac	gggactgtaa	ctcttaataa	aacggatatt	720
acttatttcag	gcgggtgatc	taccggcgct	accaaagatg	atacgttgat	taaagttgct	780
gctaattctg	acggagagga	cgttggtttc	gctaccgttc	agggttaagaa	ttatgaaatt	840
acagatgggtg	taaaaaacca	gtccactgct	gcaccaaccg	atattgctca	gaccattgat	900


```

ctggatacgg ctgatgaatt tactggggct tccactgctg atccactggc acttttagac 960
aaagctattg cacaggttga tactttccgc tcttccctcg gtgcggttca aaaccgtctg 1020
gattccgcag tcaccaacct gaacaacact actaccaacc tgtctgaagc gcagtcctcg 1080
attcaggacg cgcactatgc gaccgaagtg tccaatatgt cgaaagcgca gatcatccag 1140
caggcc

```

```

<210> 41
<211> 1506
<212> DNA
<213> Escherichia coli

```

```

<400> 41
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt ttacttctaa tattaaggc 180
ctgactcagg ctgcacgtaa cgccaatgac ggtatttctc tggcgcagac cactgaaggc 240
gcactgtctg aaatcaacaa caacttgacg cgtgtgcgtg aactgaccgt acaggcgaca 300
accggaacga actccgaatc tgacctgtcc tctatccagg acgaaatcaa atcccgctctg 360
gaagagattg accgcgtatc cggccagact cagttcaacg gcgtgaatgt gctggcaaaa 420
gacggcacca tgaaaattca ggtaggcgag aacgatgggt agactatctc tatcgatctg 480
aaaaaaatcg actcttcaac cctgggcctg accggttttg atgtttcgac gaaagcgaat 540
atttctacga cagcagtaac gggggcgcca acgaccactt atgctgatag cgcggttgca 600
attgatattg gaacggatat tagcggattt gctgctgatg ctgcgttagg aacgatcaat 660
ttcgataata caacaggcaa gtactacgca cagattacca gtgcggccaa tccgggcctt 720
gatggtgctt atgaaatcca tgtaaatgac gggatgggtt ccttctactg agcagcgagt 780
gataaacaag cgggtgctgc tccgggtact gctctgacaa gcggtaaagt tcagactgca 840
accaccacgc caggtacggc tgttgatgac actgcggcta aaactgctct ggctgcagca 900
gggtgctgaca cgagtggcct gaaactggtt caactgtcca acacggattc cgcaggtaaa 960
gtgaccaacg tgggttacgg cctgcagaat gacagcggca ctatctttgc aaccgactac 1020
gatggcacca ctgtgaccac gccgggcgca gagactgtga cttacaaaga tgcttccggt 1080
aacagcacca ctgcggctgt cacactgggt ggtctgatg gcaaaaccaa tctggttacc 1140
gccgctgacg gcaaaacgta cgggtgcgact gcaactgaat gtgctgatct gtccgatcct 1200
aataacaccg ttaaatctgt tgcagacaac gctaaaccgt tggctgccct ggatgatgca 1260
attgcgatgg tcgacaaatt ccgctcctcc ctccggtgcg tgcaaaaccg tctggattcc 1320
gcagtcacca acctgaacaa caccactacc aacctgtctg aagcgcagtc ccgtattcag 1380
gacgccgact atgcgaccga agtgccaac cgcagattat ccagcaggca 1440
ggtaactccg tgctgtccaa agctaaccag gttccgcagc aggttctgtc tctgctgcag 1500
ggtaa
1506

```

```

<210> 42
<211> 950
<212> DNA
<213> Escherichia coli

```

```

<400> 42
aacaataaac agtctgcgct gtcgacttct atcgagcgcc tctcttcttg tctgcgtatt 60
aacagcgcta aagatgacgc cgcgggccag gcgattgcta accgctttac ttctaacatc 120
aaaggctctg ctgagccgc acgtaacgcc aacgacggta tttctctggc gcagacggct 180
gaaggcgcgc tgtcagagat taacaacaac ttgcagcgta ttcgtgaact gaccgttcag 240
gcctctaccg gcacgaactc tgattccgac ctgtcttcta ttcaggacga aatcaaatcc 300
cgtcttgatg aaattgaccg tgtatctggt cagacccagt tcaacgggtg gaacgtgctg 360
tcgaaaaacg attcgatgaa gattcagatt ggtgccaatg ataaccagac gatcagcatt 420
ggcttgcaac aaatcgacag taccactttg aatctgaaag gatttaccgt gtccggcatg 480
gcggatttca gcgcggcgaa actgacgggt gctgatggta cagcaattgc tgctgcggat 540
gtcaaggatg ctgggggtaa acaagtcaat ttactgtctt acactgacac cgcgtctaac 600
agtactaaat atgcggtcgt tgattctgca accggtaaat acatggaagc cactgtagcc 660
attaccggta cggcggcgcc ggtaactgtt ggtgcagcgg aagtggcggg agccgctaca 720
gccgatccgt taaaagcact ggatgccgca atcgctaaag tcgacaaatt ccgctcctcc 780
ctcgggtgcg tttcaaacccg tctggattct gcggtcacca acctgaacaa caccaccacc 840
aacctgtctg aagcgcagtc ccgtattcag gacgccgact atgcgaccga agtgctcaac 900
atgtcgaaag cgcagattat ccagcaggcc ggtaactccg tgctggcaaa 950

```

<210> 43
 <211> 1707
 <212> DNA
 <213> Escherichia coli

<400> 43
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
 gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt ttacctctaa cattaaggt 180
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggg tcaggcttct 300
 accgggacta actccgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
 gacgaaattg accgcgtatc cgggtcaaac cagttcaacg gtgtgaacgt actggcgaaa 420
 gacggttcga tgaaaattca ggttggtgag aatgacggcc agactatcac gattgatctg 480
 aagaaaattg actcagatac gctggggctg aatgggttca acgttaatgg caaaggcact 540
 attgcgaaca aagctgctac agtcagcgat ctgaccgctg ctgggtgcaac gggaacagg 600
 ccttatgtcg tgaccacaaa caatacagca ctacagccta gcgatgcact gtctcgctg 660
 aaaaccggag atacagttac tactactggc tcgagtgtcg cgatctatac ttatgatgag 720
 gctaaaggga acttcaccac tcaagcaaca gttgcagatg gcgatgttgt taactttgag 780
 aatactctga aaccagcggc tggcactact gcacaggtg ttataactcg tagtactgg 840
 gatgtgaagt ttgatgtaga tgctaattgg gatgtgacca tcgggtgtaa agccgcgtac 900
 ctggacgcca ctggtaacct atctacaaac aaccgggca ttgcatcttc agcgaaattg 960
 tccgatctgt ttgctagcgg tagtacctta gcgacaactg gttctatcca gctgtctggc 1020
 acaacttata actttggtgc agcggcaact tctggcgtaa cctacaccaa aactgtaagc 1080
 gctgatactg tactgagcac agtgacagat gctgcaacgg ctaacacagc agttactgg 1140
 gcgacaatta agtataatca aggtattcag tctgcaacgg cgtccttcgg tgggtgta 1200
 actaatgggt ctggtaattc gaatgacacc tatactgatg cagacaaaga gctcaccaca 1260
 accgcattct acactatcaa ctacaacgtc gataaggata ccggtacagt aactgtagct 1320
 tcaaatggcg caggtgcaac tggtaaattt gcagctactg ttggggcaca ggcttatgtt 1380
 aactctacag gcaactgac cactgaaacc accagtgcag gcactgcaac caaagatcct 1440
 ctggctgccc tggatgaagc tatcagctcc atcgacaaat tccgttcac cctgggtgct 1500
 atccgaacc gtctggattc cgcggttacc aacctgaaca acaccactac caacctgtcc 1560
 gaagcgcagt cccgtattca ggacgcgcag tatgcgaccg aagtgtccaa catgtcgaaa 1620
 gcgcagatta tccagcaggc cggtaactcc gtgctggcaa aagccaacca ggtaccgcag 1680
 caggttctgt ctctgctgca ggggttaa 1707

<210> 44
 <211> 1720
 <212> DNA
 <213> Escherichia coli

<400> 44
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
 gcgaaggatg acgcgcgagg tcaggcgatt gctaaccgtt ttacttctaa tattaaggc 180
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
 gcgctgtccg aaatcaacaa caacttacag cgtgtgctg aactgacggg tcaggcgacc 300
 accggtacca actcccagtc tgatctggac tctatccagg acgaaatcaa atcccgtctg 360
 gacgaaattg accgcgtatc cggtcagact cagttcaacg gcgtgaacgt actggcaaaa 420
 gacggttcca tgaaaattca ggttggcgcg aatgatggcc agaccatcac tatcgacctg 480
 aagaagattg actcttctac gttgaaactg actgggttta acgtgaatgg ttctgggtct 540
 gtggcgata ctgcccgcac taaagacgaa ctggctgctg ctgctgcggc ggccgggtaca 600
 actcctgctg tcggtactga cggcgtgacc aaatataacc tagacgcagg gcttaacaaa 660
 gccacagcag caaacgtgtt tgcaaacctt gcagatgggt ctggtgttga tgctagcatt 720
 tccaacgggt ttggtgcagc agcagccaca gactacacct acaataaagc tacaatgat 780
 ttcaatttca atgccagcat tgctgctggt gctgcggccg gtgatagtaa cagcgcagct 840
 ctgcaatcct tcctgactcc aaaagcaggg gatacagcta acctgagcgt caaaatcggg 900
 acgacatctg ttaatgttgt tctggcgagc gatggcaaaa ttacagcgaa agatgggtca 960
 gctctgtata tcgactcaac gggtaacctg actcagaaca gcgcaggcac tgtaacagca 1020
 gcaaccctgg atggactgac caaaaacctg gatgcgacag gagctgttgg tgttgatctc 1080
 acgaccgcag atggcgcaac tatctctctg gcaggctctg ctaacgcggc aacaggctact 1140
 caatcagggt caattacact gaaaaatgtt cgtatcagtg ctgatgctct gcagtctgct 1200
 gcgaaaggta ctgttatcaa tgttgataat ggtgctgat atatttctgt tagtaaaacc 1260
 ggggtgtcgt actaccggag gtgcgcctac ttatactgat gctgatggta aattaacgac 1320
 aaccaacacc gttgattatt tctgcaaac tgatggcagc gtaaccaatg gttctggtaa 1380
 aggggtttac accgatgcag ctggtaaatc cactaccgac gctgcaacca aagccgcaac 1440
 caccaccgat ccgctgaaag cccttgatga cgcaatcagc agatcgata agttccgttc 1500

atccctgggt	gctatccaga	accgtctgga	ttccgcgggt	accaacctga	acaacaccac	1560
taccaacctg	tccgaagcgc	agtcccgtat	tcaggacgcc	gactatgcga	ccgaagtgtc	1620
caatatgtcg	aaagcgcaga	tcattccagca	ggccggtaac	tccgtgttgg	caaaagctaa	1680
ccaggtaccg	cagcagggtc	tgtctctgct	gcagggttaa			1720

<210> 45

<211> 14516

<212> DNA

<213> Escherichia coli

<400> 45

gatctgatgg	ccgtagggcg	ctacgtgctt	tctgctgata	tctgggctga	gttggaaaaa	60
actgctccag	gtgcctgggg	acgtattcaa	ctgactgatg	ctattgcaga	gttggctaaa	120
aaacagtctg	ttgatgccat	gctgatgacc	ggcgacagct	acgactgcgg	taagaagatg	180
ggctatatgc	aggcattcgt	taagtatggg	ctgcgcaacc	ttaaagaagg	ggcgaagttc	240
cgtaagagca	tcaagaagct	actgagtgag	tagagattta	cacgtctttg	tgacgataag	300
ccagaaaaaa	tagcggcagt	taacatccag	gcttctatgc	tttaagcaat	ggaatgttac	360
tgccgttttt	tatgaaaaat	gaccaataat	aacaagttaa	cctaccaagt	ttaatctgct	420
ttttgttggg	ttttttcttg	tttctggctg	catttggtaa	gacaattagc	gtgagtttta	480
gagagttttg	cgggatctcg	cggaactgct	cacatctttg	gcatttagtt	agtgcactgg	540
tagctgttaa	gccagggggc	gtagcttgcc	taattaattt	ttaacgtata	cattttattct	600
tgccgcttat	agcaaatata	gtcaatcggg	ttaaacttct	tttccattag	gtaaaagagt	660
gtttgtagtc	gctcagggaa	attggttttg	gtagtagtac	ttttcaaatt	atccattttc	720
cgatttagat	ggcagttgat	gttactatgc	tgcatataca	tcaatgtata	ttatttactt	780
ttagaatgtg	atatgaaaaa	aatagtgtac	ataggcaatg	tagcgtcaat	gatgttaagg	840
ttcaggaaag	aattaatcat	gaatttagtg	aggcaagggt	ataatgtata	ttgtctagca	900
aatgattttt	ccactgaaga	tcttaaagta	ctttcgatcat	ggggcggtta	gggggtttaa	960
ttctctctta	actcaaaggg	tattaatcct	tttaaggata	taattgctgt	ttatgaacta	1020
aaaaaaattc	ttaaggatat	ttcccagat	attgtatttt	catattttgt	aaagccagta	1080
atatttgtaa	ctattgtctc	aaagtgtgca	aaagtgccaa	ggattgttgg	aatgattgaa	1140
ggcttaggta	atgccttcac	ttattataag	ggaaaagcaga	ccacaaaaac	taaaatgata	1200
aagtggatac	aaattctttt	atataagtta	gcattaccga	tgcttgatga	tttgattcta	1260
ttaaatcatg	atgataaaaa	agatttaatc	gatcagtata	atattaaagc	taaggttaaca	1320
gtgttaggtg	ggattggatt	ggatcttaat	gagttttcat	ataaagagcc	accgaaagag	1380
aaaattacct	ttatttttat	agcaagggtta	ttaagagaga	aagggatatt	tgagtttatt	1440
gaagccgcaa	agttcgttaa	gacaacttat	ccaagttctg	aatttgtaat	tttaggaggt	1500
tttgagagta	ataatccttt	tcattacaa	aaaaatgaaa	ttgaatcgct	aagaaaagaa	1560
catgatctta	tttatcctgg	tcattgtgaa	aatgttcaag	attggttaga	gaaaagtctc	1620
gtttttgttt	tacctacatc	atatacgaga	ggcgtagcaa	gggtgatcca	agaagctatg	1680
gctattggta	gacctgtaat	aacaactaat	gtacctgggt	gtagggatat	aataaatgat	1740
ggggctcaatg	gctttttgat	acctccattt	gaaatttaatt	tactggcaga	aaaaatgaaa	1800
tattttattg	agaataaaga	taaaagtactc	gaaatggggc	ttgctggaag	gaagtttgca	1860
gaaaaaaaact	ttgatgcttt	tgaaaaaaat	aatagactag	catcaataat	aaaatcaaat	1920
aatgattttt	gacttgagca	gaaattattt	atatttcaat	ctgaaaaata	aaggctgtta	1980
ttatgaataa	agtggcatta	attactggta	tcactgggca	agatggctcc	tatttggcag	2040
aatattgttt	agaaaaagggt	tatgaagttc	atgggtattaa	acgcctgtga	tcttctttaa	2100
atactgagcg	agtggtacac	atctatcagg	attcacattt	agctaactct	aaactttttc	2160
tacactatgg	cgatttgaca	gatacttcca	atctgacctg	tattttaaaa	gaagttcaac	2220
cagatgaagt	ttacaatttg	ggggcgatga	gcatgtagc	gggtatcatt	gagtcaccag	2280
aatacactgc	tgatgttgat	gcgataggaa	cattgcgtct	tcttgaagct	atcaggatat	2340
tggggctgga	aaaaaagaca	aaattttatc	aggcttcaac	ttcagagctt	tatggtttgg	2400
ttcaagaaat	ttcacaaaaa	gagactacgc	cattttatcc	acgttcgcct	tatgctgttg	2460
caaaattata	tgcttatttg	atcactgtta	attatcgtga	gtcttatggg	atgtttgcct	2520
gcaatggtat	tctctttaac	cacgaatcac	ctcgccgtgg	cgagaccttt	gttactcgta	2580
aaataacacg	cgggtagtag	aatattgtct	aaggctctga	taaatgctta	tacttgggaa	2640
atatggattc	tctgcgtgat	tggggacatg	ctaaggatta	tgtcaaaatg	caatggatga	2700
tgctgcagca	agaaactcca	gaagattttg	taattgctac	aggaattcaa	tattctgtcc	2760
gtgagtttgt	cacaatggcg	gcagagcaag	taggcataga	gttagcattt	gaaggtgagg	2820
gagtaaatga	aaaagggtgt	gttgtttcgg	tcaatggcac	tgatgctaaa	gctgtaaacc	2880
cgggcgatgt	aattatatct	gtagatccaa	ggatattttg	gcctgcagaa	gttgaaacct	2940
tgcttggcga	tcctactaat	gcgcataaaa	aattaggtatg	gagccctgaa	attacattgc	3000
gtgaaatggt	aaaagaaatg	gtttccagcg	atttagcaat	agcgaaaaag	aacgtcttgc	3060
tgaaagctaa	taacattgcc	actaatattc	cgcaagaata	aaaaagataa	tacattaaat	3120
aattaaaaat	ggtgctagat	ttattagtag	cattattttt	ttttgggtga	ctaattgtta	3180
ttacatcaga	taaatttaga	gaaattatca	agttagttcc	attagtatca	attgatctgc	3240
taattgaaaa	cgagaatggg	gaattattat	ttggtcttag	gaataatcga	ccggccaaaa	3300
attatttttt	tgttccagggt	ggtaggattc	gcaaaaaatga	atctattaaa	aatgctttta	3360

aaagaatatac	atctatggaa	ttaggtaaag	agtatggat	ttcaggaagt	gtttttaatg	3420
gtgtatggga	acatttctat	gatgatgggt	tttttctga	aggcgaggca	acacattata	3480
tagtgctttg	ttacacactg	aaagttctta	aaagtgaatt	gaatctccca	gatgatcaac	3540
atcgtgaata	cctttggcta	actaaacacc	aaataaatgc	taaacaagat	gttcataact	3600
attcaaaaaa	ttattttttg	taatttttat	taaaaattaa	tatgcgagag	aattgtatgt	3660
ctcaatgtct	ttaccctgta	attattgccc	gaggaaccgg	aagccgtcta	tggccgttgt	3720
ctcgagtatt	ataccctaaa	caatttttaa	atttagttgg	ggattctaca	atgttgcaaa	3780
caacaattac	gcgtttggat	ggcatcgaat	gcgaaaatcc	aattgttatc	tgcaatgaag	3840
atcaccgatt	tattgtagca	gagcaattac	gacagattgg	taagctaacc	aagaatatta	3900
tacttgagcc	gaaaggccgt	aatactgcac	ctgccatagc	tttagctgct	tttatcgctc	3960
agaagaataa	tcctaattgac	gaccctttat	tattagtact	tgcggcagac	cactctataa	4020
ataatgaaaa	agcatttcga	gagtcaataa	taaaagctat	gccgtatgca	acttctggga	4080
agttagtaac	atttggatt	attccggaca	cggcaaatac	tggttatgga	tatattaaga	4140
gaagtctctc	agctgatcct	aataaagaat	tcccagcata	taatgttgcg	gagtttgtag	4200
aaaaaccaga	tgttaaaaca	gcacaggaat	atatttcgag	tgggaattat	tactggaata	4260
gcggaatggt	tttatttcgc	gccagtaaat	atcttgatga	actacggaag	tttagaccag	4320
atatttatca	tagctgtgaa	tgtgcaaccg	ctacagcaaa	tatagatatg	gactttgtcc	4380
gaattaacga	ggctgagttt	attaattgtc	ctgaagagtc	tatcgattat	gctgtgatgg	4440
aaaaaacaaa	agacgctgta	gttcttccga	tagatatgtg	ctggaatgac	gtgggttctt	4500
ggctcatcact	ttgggatata	agccaaaagg	attgccatgg	taatgtgtgc	catggggatg	4560
tgctcaatca	tgatggagaa	aatagtttta	tttactctga	gtcaagtctg	gttgcgacag	4620
tcggagtaag	taatttagta	attgtccaaa	ccaaggatgc	tgtactggtt	gcggaaccgtg	4680
ataaagtcca	aaatgttaaa	aacatagttg	acgatctaaa	aaagagaaaa	cgtgctgaat	4740
actacatgca	tcgtgcagtt	tttcgccctt	ggggtaaat	cgatgcaata	gaccaaggcg	4800
atagatatag	agtaaaaaaa	ataatagtta	aaccaggaga	agggttagat	ttaaggatgc	4860
atcatcatag	ggcagagcat	tggattgttg	tatccggtac	tgtctaaagt	tcactaggta	4920
gtgaagttaa	actattagtt	tctaagtgtg	ctatatatat	ccctcaggga	gcaaaatata	4980
gtcttgagaa	tccaggcgta	atacctttgc	atctaattga	agtaagtctt	ggtgattacc	5040
ttgaatcaga	tgatatagtg	cgttttactg	acagatataa	cagtaaacaa	ttcctaagac	5100
gagattgata	aatatgaata	aaataacttg	cttcaaagca	tatgatatac	gtgggcgtct	5160
tggtgctgaa	tgaatgatg	aaatagcata	tgaatttgtt	cgcgcttatg	gtgagttttt	5220
taaacctcaa	actgtagttg	tgggaggaga	tgtctcgctt	acaagtgaga	gtttaaagaa	5280
atcactctca	aatgggctat	gtgatgcagg	cgtaaatgtc	ttagatcttg	gaatgtgtgg	5340
tactgaagag	atatattttt	ccacttggtg	tttaggaatt	gatggtggaa	tcgaggtaac	5400
tgaagccat	aatccaattg	attataatgg	aatgaaatta	gtaaccaaag	gtgctcgacc	5460
aatcagcagt	gacacaggtc	tcaaagatat	acaacaatta	gtagagagta	ataattttga	5520
agagctcaac	ctagaaaaaa	aagggaatat	taccacaaat	tccacccgag	atgcctacat	5580
aaatcatttg	atgggctatg	ctaactctga	aaaaataaaa	aaaatcaaaa	tagttgtgaa	5640
ttctgggaat	ggtgcagctg	gtcctgttat	tgatgctatt	gaggaatgct	ttttacggaa	5700
caatatcccg	attcagtttg	taaaaaataa	taatacaccc	gatggtaatt	ttccacatgg	5760
tatccctaata	ccattactac	ctgagtgcag	agaagatacc	agcagtgccg	ttataagaca	5820
tagtgctgat	tttggtattg	catttgatgg	tgattttgat	agggtgtttt	tctttgatga	5880
aaatggacaa	tttattgaag	gatactacat	tgttggttta	ttagcggaag	tttttttagg	5940
gaaatatcca	aacgcaaaaa	tcattcatga	tcctcgccct	atatggaata	ctattgatat	6000
cgtagaaaagt	catgggtggt	tacctataat	gactaaaacc	ggctcatgct	acattaagca	6060
aagaatgcgt	gaagaggatg	ccgtatatgg	cgcgcgaatg	agtgcgcac	attattttta	6120
agatttttga	tactgcgata	gtggaatgat	tccttggtat	tttaatttgt	aacttttgag	6180
tctgacaaat	aaaaaattag	gtgaactggt	ttgtggttgt	ataaacgact	ggccggcaag	6240
tggagaaaata	aactgtacac	tagacaatcc	gcacaaatga	atagataaat	tatttaatcg	6300
ttacaaagat	agtgccttag	ctgttgatta	cactgatgga	ttactatagg	agttctctga	6360
ttggcgtttt	aatgtttag	gctcaaatac	agaacctgta	gtacgattga	atgtagaatc	6420
taggaataat	gctattctta	tgcaggaaaa	aacagaagaa	attctgaatt	ttatatcaaa	6480
ataaattttg	acctgagttc	ataatgggaa	caagaaatat	atgaaagtac	ttctgactgg	6540
ctcaactggc	atggttggtg	agaatatatt	agagcatgat	agtgcgaagt	aatataatat	6600
acttactcca	accagctctg	atttgaattt	attagataaa	aatgaaatag	aaaaattcat	6660
gcttatcaac	atgccagact	gtattataca	tgcagcggga	ttagttggag	gcattcatgc	6720
aaatataagc	aggccgtttg	attttctgga	aaaaaatttg	cagatggggt	taaaatttagt	6780
ttccgtcgca	aaaaaactag	gtatcaagaa	agtgcctaac	ttgggtagtt	catgcatgta	6840
ccccaaaaac	tttgaagagg	ctattcctga	gaaagctctg	ttactgggtg	agctagaaga	6900
aactaatgag	ggatatgcta	ttgcgaaaaa	tgctgtagca	aaagcatgcg	aatatatatc	6960
aagagaaaaa	tctaattatt	tttataaaac	aattatocca	tgtaatttat	atgggaaata	7020
tgataaaatt	gatgataact	cgtcacatat	gattccggca	gttataaaaa	aaatccatca	7080
tgcgaaaaatt	aataatgtcc	cagagatcga	aatttggggg	gatggtaatt	cgcgccgtga	7140
gtttatgtat	gcagaagatt	tagctgatct	tattttttat	gttattccta	aatagaattt	7200
catgcctaata	atggtaaatg	ctggtttagg	ttacgattat	tcaattaatg	actattataa	7260
gataatttga	gaagaaattg	gttatactgg	gtatttttct	catgatttaa	caaaaccaac	7320
aggaatgaaa	cgggaagctag	tagatatctc	attgcttaatt	aaaattgggt	ggtcaagtca	7380

ctttgaactc	agagatggca	tcagaaagac	ctataattat	tacttggaga	atcaaaaata	7440
atgattacat	accacttgc	tagtaatact	tgggatgaat	atgagtatgc	agcaatacag	7500
tcagtaattg	actcaaaaat	gtttaccatg	ggtaaaaagg	ttgagttata	tgagaaaaat	7560
tttgctgatt	tgtttggtag	caaatatgcc	gtaatggtta	gctctgggtc	tacagctaata	7620
ctgttaatga	ttgctgccct	tttcttctac	aataaaacaa	aacttaaaaag	aggtgatgaa	7680
ataatagtag	ctgcagtgtc	atgggtctacg	acataattacc	ctctgcaaca	gtatggctta	7740
aagggtgaagt	ttgtcgatat	caataaagaa	acttttaata	ttgatatcga	tagtttgaaa	7800
aatgctattt	cagataaaaac	aaaagcaata	ttgacagtaa	atttatttagg	taatcctaata	7860
gattttgcaa	aaataaatga	gataataaat	aatagggata	ttatcttact	agaagataac	7920
tgtgagtcga	tgggcgcggg	ctttcaaaaat	aagcaggcag	gcacattcgg	agttatgggt	7980
accttttagtt	ctttttactc	tcatacatata	gctacaatgg	aagggggctg	cgtagttact	8040
gatgatgaag	agctgtatca	tgtattgttg	tgccttcgag	ctcatgggtg	gacaagaaat	8100
ttaccaaaaag	agaatatggg	tacaggcact	aagagtgtatg	atattttcga	agagtgcgtt	8160
aagtttggtt	taccaggata	caatgttcgc	ccacttgaaa	tgagtgggtgc	tattgggata	8220
gagcaactta	aaaagttacc	aggtttttata	tccaccagac	gttccaatgc	acaataattt	8280
gtagataaat	ttaaagatca	tccattccctt	gatatacaaaa	aagaagttgg	tgaaagtagc	8340
tggtttggtt	tttcttctcg	tataaaggag	ggagctgcta	ttgagaggaa	gagtttagta	8400
aataatctga	tctcagcagg	cattgaatgc	cgaccaattg	ttactgggaa	ttttctcaaa	8460
aatgaacgtg	ttttgagtta	ttttgattac	tctgtacatg	atacggtagc	aaatgccgaa	8520
tatatagata	agaatgggtt	ttttgtcggg	aaccaccaga	tacctttgtt	taatgaaata	8580
gattatctac	gaaaagtatt	aaaaataacta	acgaggcact	ctattttcgaa	tagagtgcct	8640
ttaaagatgg	attacagtg	aaaaaaat	tagcgtttgg	ctatttctaaa	gtactaccac	8700
cgggtattga	acagtttgct	aatccaat	gcattcttc	tatcacacca	ctaatactca	8760
accacctggg	taagcaaagc	tatggtaatt	ggattttatt	aattactatt	gtatcttttt	8820
ctcagttaat	atgtggagga	tgttccgc	ggattgcaaa	aatcattgca	gaacagagaa	8880
ttcttagtga	tttatcaaaa	aaaaatgctt	tacgtcaaat	ttcctataat	ttttcatttg	8940
ttattatcgc	atlttgcggt	ttgatttctt	ttcttatatt	aagtatttgt	ttcttcgatg	9000
ttgcgaggaa	taattcttca	ttcttattcg	cgattattat	ttgtgggttt	tttcagggaag	9060
ttgataat	atlttagtgg	gcgtcaaaaag	gttttgaaaa	atttaagtta	tcattgtttt	9120
ttgaagtaat	tacaagagtg	ctctgggctt	ctatagtaat	atatggcatt	tacggaaatg	9180
cactcttata	ttttacatgt	ttagccttta	ccattaaagg	tatgctaaaa	tatatctctg	9240
tatgtctgaa	tattaccggt	tgtttcatca	atcctaattt	taatagagtt	gggattgtta	9300
atlttgtaaa	tgagtcaaaa	tggtgtttc	ttcaattaac	tggtggcgct	tcacttagtt	9360
tgtttgatag	gctcgtataa	ccattgattt	tatctgtcag	taaactggct	tcttatgtcc	9420
cttgccttca	actagctcaa	ttgatgttca	ctctttctgc	gtctgcaaat	caaatactac	9480
taccaatgtt	tgctagaatg	aaagcatcta	acacatttcc	ctctaattgt	tttttcaaaa	9540
ttctgcttgt	atcactaatt	tctgttttgc	cttgtcttgc	gttattcttt	tttggctcgtg	9600
atatattatc	aatatggata	aaccctacat	ttgcaactga	aaattataaa	ttaatgcaaa	9660
ttttagctat	aagttacatt	ttattgtcaa	tgatgacatc	ttttcatttc	ttgttattag	9720
gaattggtaa	atctaagctt	gttgcaaat	taaatctggg	tgcaaggctc	gcacttgctg	9780
cttcaacgtt	atacgcagct	cattatggcc	ttatgcaat	atctatggta	aaaataattt	9840
atccggcttt	tcaattttat	tacctttatg	tagcttttgt	ctattttaat	agagcgaaaa	9900
atgtctattg	atlttacttt	ttcaattact	gaaatcgcaa	ttgttttttc	ttgcactatt	9960
tacataattt	ctcaatgttt	gttaatgcgg	aggatctatt	tagataaaaag	tatttttaatt	10020
cttttatgct	tgctcttttt	tttagtaatc	attcaacttc	ctgagcttaa	tgtaaacggt	10080
ttggctgatt	ctttaaagt	atcactgcct	ttattgtatg	tctttatcgc	ttttcaaaa	10140
ccgaaattat	gcttgtgggt	tattattgca	ttgttgtttt	tgaactctgc	atlttaatttt	10200
ttatatatta	agacattcga	taagtttagc	tcatttcctt	ttactttttt	tattattgctg	10260
ttttacttgt	ttagattggg	aattggtaat	ttaccgggtt	ataaaaaata	aaaattttac	10320
gcgttgattt	ttctctttat	attaatagac	ataatgcagt	cattgttaat	aaattatagg	10380
gggcagattt	tatatccgt	aatttgcac	ctgataactg	tgtttaaaagt	taatttaaga	10440
aaaaagattc	catacttttt	tttaatgctg	ccagttttat	atgtaattat	tatggcttat	10500
attgggtttt	attatttcaa	taaaggcgta	actttttttg	aacctacagc	aagtaatat	10560
gaacgtacgg	ggatgatata	ttatttggtt	tcacagcttg	gtgattatat	attccatggg	10620
atggggacat	taaatttctt	aaataacggc	ggacaatata	agacgttata	tggaactcca	10680
tcattaaattc	ctaatagacc	tcattgattt	ttattacggg	tctttataag	tattgggtgtg	10740
ataggagcat	tggtttatca	ttctatat	tttgtttttt	ttaggagaat	atctttctta	10800
ttatatgaga	gaaatgctcc	tttcattgtt	gtaagtgtgt	tgttactgtt	acaagttgtg	10860
ttaatattata	cattaaaccc	ttttgatgct	tttaatcgat	tgatttgcgg	gcttacagtt	10920
ggagttgttt	atggatttgc	aaaaattaga	taagtatacc	tgtaatggaa	atlttagacg	10980
tccacttggt	tcaataatca	ttgcaactta	ttaattctgaa	cttgatatag	ctaagtggtt	11040
gcaatcggt	actaatcaat	cttataagaa	tattgaaatc	ataataatgg	atggaggatc	11100
ttctgataaa	acgcttgata	ttgcaaaatc	gtttaaagac	gaccgaataa	aaatagtttc	11160
agagaaagat	cgtggaattt	atgatgcctg	gaataaagca	gttgatttat	ccattgggtga	11220
ttgggtagca	tttattgggt	cagatgatgt	tactatcat	acagatgcaa	ttgcttcat	11280
gatgaagggg	gttatgggt	ctaattggcg	ccctgtgggt	tatggggagg	cagcgacga	11340
aggtcccgat	aggaacatat	ctggattttc	aggcagtgaa	tggtacaacc	taacaggatt	11400

taagttta	at	tattacaa	at	gtaattt	acc	attgccc	att	atgagcg	caa	tatatct	ctg	11460
tgatttct	tc	agaacga	ac	gttttg	atat	taaatt	aaaa	attgttg	ctg	acgctg	attg	11520
gtttctg	aga	tgtttca	tca	aatgg	agtaa	agaga	agtca	ccttatt	ttta	ttaatg	acac	11580
gaccctat	t	gttaga	atgg	gatatg	gtg	ggtttc	gact	gatatt	ctt	ctcaag	ttaa	11640
aactacg	cta	gaaagtt	tca	ttgtac	gcaa	aaaga	ataat	atatc	ctgtt	taaaca	taca	11700
gctgatt	ctt	agatat	gcta	aaatt	ctggt	gatggt	tagcg	atcaaaa	ata	tttttg	gcaa	11760
taatgtt	at	aaatta	atgc	ataac	gggta	tcatt	cccta	aagaaa	atca	agaata	aaaat	11820
atgaag	attg	tttatata	aat	aacc	gggctt	acttg	tgtgtg	gagcc	gaaca	ccttat	gacg	11880
cagttag	cag	accaa	atggt	tatac	gcggg	catgat	gtta	atatt	atttg	tcta	actggt	11940
atatctg	agg	taaag	ccaac	acaaa	atatt	aatatt	catt	atgtta	aatat	ggata	aaaaa	12000
tttaga	agct	tttttag	agc	tttatt	tcaa	gtaaaa	aaaaa	taatt	gtcgc	cttaa	agcca	12060
gatata	aatac	atagtc	atat	gtttc	atgct	aatatt	tttta	gtcgt	ttttat	taggat	gctg	12120
attccag	cgg	tgccct	gat	atgtac	cgc	caca	acaaaa	atga	aggtg	caatg	caagg	12180
atgttt	tgt	atcgact	gag	tgatt	tttta	gctt	ctatta	ctaca	aatgt	aagta	aaag	12240
gctgtt	caag	agtttat	agc	aagaa	aggct	acac	ctaaaa	ataaa	atagt	agag	attccg	12300
aatttt	atta	ataca	aataa	atttg	atttt	gatatta	atg	tcaga	aaaga	aacgc	gagat	12360
gctttta	aat	tgaa	agacag	tacag	cagta	ctgct	cgcag	tagga	agact	tgtt	gaagca	12420
aaagact	atc	cgaac	ttatt	aaatg	caata	aatc	atttga	ttctt	tcaaa	aacat	caaat	12480
tgta	atgatt	ttatt	tttgc	tattg	ctggc	gatgg	cgcat	taaga	aaata	attatt	ggat	12540
ttgg	tttgc	aattg	aatct	tgtg	gataa	gtttt	cttct	tgggg	caaag	aagt	gatatt	12600
aaaga	attaa	tgtgt	gctgc	agat	ctttt	gtttt	gagtt	ctgag	tggga	agg	ttttggt	12660
ctcgt	tgttg	caga	agctat	ggcgt	gtgaa	cgtcc	ggtt	ttgct	accga	ttctg	gttga	12720
gttaa	agaag	tcgtt	ggacc	tcata	atgat	gttat	ccctg	tcag	taatca	tatt	ctgttg	12780
gcag	agaaaa	tcgct	gagac	actt	aaaaa	gatg	ataacg	caaga	aaaaa	aatag	gtatg	12840
aaaa	atagag	aatata	attgt	ttcca	atttt	tcaatt	aaaaa	cgat	agttag	tgagt	gggag	12900
cgctt	atatt	ttaaa	tattc	caag	cgta	aatata	atttg	attg	aaaaa	taagt	ttgt	12960
ctctg	gatgc	aatagt	tttct	ctatg	ctgtt	ttttt	acttg	ctccg	tattt	ttact	tatat	13020
ctgg	atttg	ttata	tatca	gtatta	atct	gtct	caactt	cat	ctagact	acatt	caagc	13080
cgcg	catgcg	tcgcg	cgggtg	actac	acctg	acagg	agtat	gta	atgtcca	agca	acagat	13140
ggcgt	cgtc	ggtat	ggcag	tgat	ggggc	caac	ctggc	ctca	acatcg	aaag	cgcgcg	13200
ttata	cgcgt	tccat	cttca	accg	ctccc	cgag	aaaaa	gaaga	agttg	ttgc	cagaa	13260
cccgg	ataag	aaact	ggttc	cttatt	tacac	ggtg	aaagag	ttcgt	cgagt	ctctt	gaaac	13320
cccac	gctcgt	atcct	gttaa	tggt	aaaagc	agggg	cggga	actg	atgctg	ctat	cgattc	13380
cctga	agccg	tatct	ggata	aagg	cgcacat	catt	attgat	ggtg	gcaaca	cctt	cttcca	13440
ggac	actatc	cgtc	gtaacc	gtga	actgtc	cgcg	gaaggc	ttta	acttca	tcgg	taccgg	13500
cgtgt	ccggc	ggtg	aaagag	gcgc	ccctgaa	agg	cccatct	atcat	gccag	gtgg	ccagaa	13560
aga	agcgtat	gagct	ggttg	cgct	atcct	gacca	agatt	gctg	cgggtg	ctga	agatgg	13620
cga	accatgt	ataact	taca	tcggt	gctga	cggt	gcgggt	cact	acgtga	agat	ggtgca	13680
caac	ggtatc	gaat	atggcg	atat	gcagct	gatt	gctgaa	gcct	atttctc	tgct	taaagg	13740
cggc	ctta	at	ctctaacg	aagag	ctggc	aacc	actttt	accg	agtgga	atga	aggcga	13800
gcta	agtagc	tacct	gattg	acat	cacc	aa	acattctt	acc	aaaaa	atga	agagg	13860
taa	acactg	gttg	atgtga	tcct	ggacga	agct	gcgaac	aaag	gcaccg	gtaa	atggac	13920
cag	ccagagc	tctct	ggatc	tggg	tgaacc	gctgt	cgcgtg	atc	accgaat	ccgt	attcgc	13980
tcg	ctacatc	tctt	ctctga	aag	accagcg	catt	gcggca	tcta	aaagtgc	tgt	ctggtcc	14040
gcagg	ctaaa	ctgg	ctgggtg	ataa	agcaga	gtt	cgttgag	aaag	tccgctc	gcgc	gctgta	14100
cct	gggtaaa	atcgt	ctctt	atg	cccaagg	ctt	cttca	ctgc	gtgccc	cgct	gtacga	14160
ata	caactgg	gat	ctgaact	acgg	cgaat	cg	gaagatc	ttcc	gcgcg	gct	gcatcat	14220
tcgt	gcgcag	ttc	ctgcaga	aaatt	actga	cg	ctatgct	gaaa	acaaag	gcatt	gctaa	14280
cct	gttgctg	gct	ccgtact	tca	aaaat	at	cgctgat	tat	cagcaag	cgct	gctgta	14340
tgt	agtggct	tat	gctgtgc	aga	acgggt	at	ccgtaccg	ac	cttctctg	cag	cggtagc	14400
ctact	accagc	agct	accgtt	ctgc	gggtact	gc	gggtat	ctg	attcagg	cac	agcgtga	14460
ttact	tcggt	gcgc	acacgt	ataa	acgcac	tg	ataaagaa	ggt	gtgttcc	ac	accg	14516

<210> 46

<211> 1380

<212> DNA

<213> Escherichia coli

<400> 46

aacaaa	atctc	agtctt	ctct	tagct	ctgct	attgag	cgctc	tgtctt	cttg	tctgc	gtatt	60	
aacag	cgcaa	aagac	gatgc	agcag	gtcag	gcgatt	gcta	accgtt	tttac	ggcaa	atatt	120	
aaagg	tctga	cccag	gcttc	ccgta	acgcg	aatgat	ggta	tttct	gttgc	gcag	accact	180	
gaagg	tgcgc	tgaat	gaaat	taaca	acaac	ctgc	agcgt	ttcgt	gaact	ttct	gttcag	240	
gcaact	aacg	gtact	aactc	tgac	agcgt	cttt	cttcta	tccag	gctga	aatt	actcaa	300	
cgtct	ggaag	aaatt	gaccg	tgtat	ctgag	caaact	cagt	ttac	cggcgt	gaaag	tcctt	360	
gctg	aaaata	atg	aaatgaa	aatt	caggtt	ggtg	cta	atg	gtgaa	aac	catc	atc	420
aatct	ggcaa	aaatt	gatgc	gaaa	actctc	ggct	ggacg	gttt	tta	atat	cgat	ggcgcg	480

```

cagaaagcaa cccgcagtgga cctgatttct aaattttaaag cgacagggtac tgataattat 540
caaattaacg gtactgataa ctatactgtt aatgtagata gtggagtagt acaggataaa 600
gatggcaaac aagttttatgt gagtgctgcg gatgggtcac ttacgaccag cagtgtact 660
caattcaaga ttgatgcaac taagcttgca gtggctgcta aagatttagc tcaaggtaat 720
aagattgtct acgaagggtat cgaatttaca aataccggca ctggcgctat acctgccaca 780
ggtaatgggt aattaaccgc caatgttgat ggtaaggctg ttgaattcac tatttcgggg 840
agtgtgata catcagggtac tagtgcaacc gttgccccta cgacagccct atacaaaaat 900
agtgcagggc aattgactgc aacaaaagtt gaaaataaag cagcgacact atctgatctt 960
gatctgaacg ctgccaagaa aacaggaagc acgttagttg ttaacggtgc aacttacgat 1020
gttagtgtag atggtaaaac gataacggag actgcttctg gtaacaataa agtcatgtat 1080
ctgagcaaat cagaagggtg tagcccgatt ctggtaaacg aagatgcagc aaaatcgttg 1140
caatctacca ccaacccgct cgaaactatc gacaaagcat tggctaaagt tgacaatctg 1200
cgttctgacc tcggtgcagt acaaaaccgt ttcgactctg ccatcaccaa ccttggcaac 1260
accgtaaaaca acctgtcttc tgcccgtagc cgtatcgaag atgctgacta cgcgaccgaa 1320
gtgtctaaca tgtctcgtgc gcagatcctg caacaagcgg gtacctctgt tctggcacag 1380

```

```

<210> 47
<211> 1497
<212> DNA
<213> Escherichia coli

```

```

<400> 47
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtcctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcgcgagg tcaggcgatt gctaaccggt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtctg aaatcaacaa caacttacag cgtgtgcgtg aactgacggt acaggccact 300
accggtacta actctgagtc tgatctgtct tctatccagg acgaaattaa atcccgtctg 360
gatgaaattg accgcgtatc tggtcagacc cagttcaacg gcgtgaacgt gctggcaaaa 420
aatggctcca tgaaaaatcca ggttgccgca aatgataacc agactatcac tatcgatctg 480
aagcagattg atgctaaaac tcttggcctt gatggtttta gcgttaaaaa taacgatata 540
gttaccacta gtgctccagt aactgctttt ggtgctacca ccacaaacaa tattaacctt 600
actggaatta cctttctac ggaagcagcc actgatactg gcggaactaa cccagcttca 660
attgaggggtg ttataactga taatggtaat gattactatg cgaaaatcac cgggtgtgat 720
aacgatggga agtattacgc agtaacagtt gctaattgat gtacagtgc aatggcgact 780
ggagcaacgg caaatgcaac tgtaactgat gcaataacta ctaaagctac aactatcact 840
tcaggcggtg cacctgttca gattgataat actgcagggt ccgcaactgc caaccttgg 900
gctgttagct tagtaaaact gcaggattcc aagggtaatg ataccgatac atatgcgctt 960
aaagatacaa atggcaatct ttacgctgcg gatgtgaatg aaactactgg tgctgtttct 1020
gttaaaacta ttacctatac tgactcttcc ggtgcccga gttctccaac cgcggtcaaa 1080
ctgggcccgg atgatggcaa aacagaagtg gtcgatattg atggtaaaac atacgattct 1140
gccgatttaa atggcggtaa tctgcaaaca ggtttgactg ctggtgggtg ggctctgact 1200
gctgttgcaa atggtaaaac cacggatccg ctgaaagcgc tggacgatgc tatcgcatct 1260
gtagacaaat tccgttcttc cctcggtgcg gtgcaaaacc gtctggattc cgcggttacc 1320
aacctgaaca acaccactac caacctgtct gaagcgcagt cccgtattca ggagccgac 1380
tatgcgaccg aagtgtccaa tatgtcgaaa gcgcagatca tccagcaggc cggttaactcc 1440
gtgttggaag aagctaacca ggtaccgcag caggttctgt ctctgctgca ggggttaa 1497

```

```

<210> 48
<211> 1695
<212> DNA
<213> Escherichia coli

```

```

<400> 48
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtcctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcgcgagg tcaggcgatt gctaaccggt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtctg aaatcaacaa caacttacag cgtattcgtg aactgacggt tcaggcttct 300
accgggaacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc cggtcacacc cagttcaacg gtgtgaacgt actggcgaaa 420
gacggttcga tgaaaattca ggttgggtcg aatgacggcc agactatcac tattgatctg 480
aagaaaattg actctgatac gctggggctg aatggtttta acgttaacgg caaagggtact 540
attgcgaaca aagcggcaac cattaagtat ctggcggcga cggggcgcaa tgttactaac 600
tcaagcaata ttgttgcac gacaaagttc aatgccttgg atgcagcgac tgcatttagc 660
aaactcaaag atggtgattc tgttgcggtt gctgctcaga aatatactta taacgcctcg 720

```

accaatgatt	ttacgacaga	aaatacagta	gcgacaggca	ctgcaacgac	agatcttggc	780
gctactctga	aggctgctgc	tgggcagagt	caatcaggta	catatacctt	tgcaaatggg	840
aaagttaact	ttgatgttga	tgcaagcggg	aatatcacta	ttggcggcga	aaaggctttc	900
ttggttggtg	gagcgctgac	tactaacgat	cccaccggct	ccactccagc	aacgatgtct	960
tccttgttta	aggcgcggga	tgacaaagat	gccgctcaat	cctcgattga	ttttggcggg	1020
aaaaaatacg	aatttgctgg	tggcaattct	actaatggtg	gcggcgttaa	attcaagac	1080
acggtgtctt	ctgacgcgct	tttggtctag	gttaaagcgg	atagtactgc	taataatgta	1140
aaaatcacct	ttaacaatgg	tcctctgtca	ttcactgcat	cgttccaaaa	tggtgtatct	1200
ggctccggcg	catcgaatgc	agcctacatt	gatagcgaag	gcgaactgac	aactactgaa	1260
tcctacaaca	caaattattc	cgtagacaaa	gacacggggg	ctgtaagtgt	tacagggggg	1320
agcgggtacg	gtaaatacgc	cgcaaacgtg	ggtgctcagg	cttatgtagg	tgcagatggg	1380
aaattaacca	cgaatactac	tagtaccggc	tctgcaacca	aagatccact	aaatgcgctg	1440
gatgaggcaa	ttgcatccat	cgacaaattc	cgttcttccc	tgggggctat	ccagAACCGT	1500
ctggattccg	cagtcaacca	cctgaacaac	accactacca	acctgtctga	agcgcagtc	1560
cgtattcagg	acgccgacta	tgcgaccgaa	gtgtccaaca	tgtcgaaagc	gcagatcatc	1620
cagcaggccg	gtaactccgt	gttggcaaaa	gctaaccagg	taccgcagca	ggttctgtct	1680
ctgctgcagg	gttaa					1695

<210> 49
 <211> 1164
 <212> DNA
 <213> Escherichia coli

<400> 49						
aacaagaacc	agtctgcgct	gtcaggttct	atcgagcgtc	tgtcttctgg	cttgcgattt	60
aacagcgcga	aggatgacgc	cgcgggtcag	gcgattgcta	accgtttttac	ttctaactatt	120
aaaggcctga	ctcaggtctg	acgtaacgcc	aacgacggta	tttctgttgc	gcagaccacc	180
gaaggcgcgc	tgtccgaaat	taacaacaac	ttacagcgctg	tgcgtgagct	gactgttcag	240
gcgaccaccg	gtactaactc	tgagtctgac	ctgtcttcta	tccaggacga	aatcaaatct	300
cgcctgggaag	agattgatcg	gttttcaagt	cagactcaat	ttaacggcgt	gaatgttttg	360
gctaaagatg	ggaaaatgaa	cattcagggt	ggggcaagtg	atggacagac	tatcactatt	420
gatctgaaaa	agatcgattc	atctacacta	aaactctcca	gttttgatgc	tacaaacttg	480
ggcaccagtg	ttaaagatgg	ggcaccatc	aataagcaag	tggcagtaga	tgctggcgac	540
tttaaagata	aagcttcagg	atcgtaggt	accctaaaat	tagttgagaa	agacggtaag	600
tactatgtaa	atgacactaa	aagtagtaag	tactacgatg	ccgaagtaga	tactagtaag	660
ggtgaaatta	acttcaactc	tacaaatgaa	agtggaaacta	ctcctactgc	agcgacggaa	720
gtaactactg	ttggccgcga	tgtaaaattg	gatgcttctg	cacttaaagc	caaccaatcg	780
cttgctcgtg	ataaagataa	aagcggcaat	gatgcttata	tcattcagac	caaagatgta	840
acaactaatc	aatcaacttt	caatgccgct	aatatcagtg	atgctgggtg	tttatctatt	900
ggtgcattcta	caaccgcgcc	aagcaattta	acagctgacc	cgtttaaggc	tcttgatgat	960
gcaattgcat	ctgttgataa	attccgctct	tctctcggtg	ccgttcagaa	ccgtctggat	1020
tctgccattg	ccaacctgaa	caacaccact	accaacctgt	ctgaagcgca	gtcccgtatt	1080
caggacgctg	actatgcgac	cgaagtgtcc	aacatgtcga	aagcgcagat	tatccagcag	1140
gccggtaact	ccgtgctggc	aaaa				1164

<210> 50
 <211> 1818
 <212> DNA
 <213> Escherichia coli

<400> 50						
atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagtctg	cgtgtctgag	ttctatcgag	cgtctgtctt	ctggettgcg	tattaacagc	120
gcgaaggatg	acgcagcggg	tcaggcgatt	gctaaccggt	tcacctctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgttaacgat	ggtatctctc	tggcgcagac	cactgaaggc	240
gcactgtctg	agattaacaa	caacttacaa	cgtgtgcgctg	agttgactgt	acaggcgacc	300
accggtacta	actctgattc	tgacctggct	tctatttcagg	acgaaatcaa	atcccgtttg	360
tctgaaattg	accgcgtatc	cgggcagacc	cagttcaacg	gcgtgaacgt	attgtctaaa	420
gatggctccc	tgaaaattca	ggttggcgca	aatgatggtc	agactatctc	tatcgacctg	480
aagaaaattg	actctgatac	tctgggtttg	aatggtttca	acgttaatgg	ttctggtagc	540
attgcaaaca	aagcggccac	aatcagtgac	ttgactgctc	agaaagccgt	tgacaacggg	600
aatggtactt	ataaagttac	aactagcaac	gctgcactta	ctgcatctca	ggcattaagt	660
aagctgagtg	atggcgatag	tgtagatatt	gcaacctatg	ctggtgggtac	aagttcaaca	720
gttagttata	aatacgacgc	agatgcaggt	aacttcagtt	ataacaatac	tgcaaacaaa	780
acaagtgcctg	cggctggaac	tctggcagat	actcttctcc	cggcagctgg	ccagactaaa	840
accggtactt	acaaggctgc	tactgggtgat	gttaacttta	atgttgacgc	aactggtaat	900


```

ctgacaattg gcggacagca agcctacctg actactgatg gtaaccttac aacaaacaac 960
tccggtgggtg cggctactgc aactcttaaa gagctgttta ctcttgctgg cgatggtaaa 1020
tctctgggga acggcggtac tgctaccgtt actctggata atactacgta taatttcaaa 1080
gctgctgcga acgttactga tgggtgtggg gtcctcgctg ctgctgggtg aacttataca 1140
gccactgttt ctaaaagatgt cattctggca caactgcaat ctgcaagtca ggcagcagca 1200
accgctaccg acggtgatac tgtcgcaacg atcaactata aatctgggtg catgatcggt 1260
tccgctacct ttaccaatgg taaaggctact gccgatggta tgacttctgg tacaactcca 1320
gtcgtagcta caggtgctaa agctgtatat gttgatggca acaatgaact gacttccact 1380
gcatcttacg atacgactta ctctgtcaac gcagatacag gcgcagtaaa agtgggtatca 1440
ggtaactggta ctggtaaatt tgaagctgtt gctgggtgcg atgcttatgt aagcaaaagat 1500
ggcaaatata cgacagaaac caccagtgcg gctactgcaa ccaaagatcc tttggctgcc 1560
ctggatgctg ctatcagctc catcgacaaa ttccgttcct ccctgggtgc tatccagaac 1620
cgtctggatt ccgcagtcac caacctgaac aacaccacta ctaacctgtc tgaagcgcag 1680
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca atatgtcgaa agcgcagatc 1740
atccagcagg ccggtaaact tgtgttggca aaagctaacc aggtaccgca gcaggttctg 1800
tctctgctgc aggggttaa

```

```

<210> 51
<211> 1344
<212> DNA
<213> Escherichia coli

```

```

<400> 51
atggcacaaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaacggtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggt tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
gacgaaattg accgcgtttc cggtcagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacggttcga tgaagattca ggttggcgcg aatgacgggc agaccatctc tatcgatttg 480
cagaaaattg attcttcaac gctgggattg aaagggttct cggtatcagg gaacgcatta 540
aaagttagcg atgcgataac tacagtctct ggtgctaata ctggcgatgc cccggttacg 600
gttaaatttg gtgcgaacga taccgtctgt gccgcaatgg ctaaaacatt gggaataagt 660
gatacatcag gcttgtccct acataacgta caaagcgcgg atggtaaagc gacaggaacc 720
tatgttgttc aatctggtta tgacttctat tcggttccg ttaatgctgg tggcgttctt 780
acgcttaata ccaccaatgt tacttttact gatcctgcga acggtgttac cacagcaaca 840
cagacaggtc agcctatcaa ggtcacgacg aatagtctg gcgcggctgt tggctatggt 900
actattcaag gcaaagatta ccttgcctgg gcagacggta aggatgcaat tgaaaacggt 960
ggtgacgctg caacaaatga agacacaaaa atccaactta ccgatgaact cgatgttgat 1020
ggttctgtaa aaacagcggc aacagcaaca ttttctggta ctgcaaccaa cgatccgctg 1080
gcacttttag acaaagctat ctgcgaagtt gatactttcc gtcctccct cggtgccgta 1140
caaaaccgtc tggattctgc ggtcaccaac ctgaataaca ccaccaccaa cctgtctgaa 1200
gcgcagtcct gtattcagga cgccgactat gcgaccgaag tgtccaacat gtcgaaagcg 1260
cagatcatcc agcaggcggg taactctgtg ctgtctaaag ctaaccaggt accgcagcag 1320
gttctgtctc tgctgcaggg ttaa

```

```

<210> 52
<211> 2599
<212> DNA
<213> Escherichia coli

```

```

<400> 52
cttctcttag ctctgctatt gagegtctgt cttctgggtc gcgtattaac agcgcaaaag 60
acgatgcagc aggtcaggcg attgctaacc gttttacggc aaatattaaa ggtctgacct 120
aggcttcccg taacgcgaat gatggtatct ctggttgcga gaccactgaa ggtgcgctga 180
atgaaattaa caacaacctg cagcgatttc gtgaactttc tgttcaggca actaacggta 240
ctaactctga cagcgatctt tcttctatcc aggctgaaat tactcaacgt ctggaagaaa 300
ttgaccgtgt atctgagcaa actcagttta acggcgtgaa agtccttgct gaaaaataatg 360
aaatgaaaat tcaggttggg gctaatagat gtgaaaccaa tgacctgccc ccacgattag 420
atacaacact cagttagtaa cgtcggaatc ttcatctca gaatgacctt ttctccagcg 480
cgctgcaaat tcagacgggt tctgataatt cagcgtggag tgcgggcggc attcgttata 540
atcctgccgc cagtcattaa taattttcct ggcatagaac atatcgctga accagtgtc 600
attcaaacat tcatacgcaa atcgtccgtt aaagctctca ataatccgt tctgcgttgg 660
cttgccccgc tggattaagc gcaactcaac accatgctca aaggccatt gatccagtgc 720
acggcaagtg aactccggcc cctggtcagt tcttatcgta gccggatagc ctcgaaacag 780
tgcaatgctg tccagaatac gcgtgacctg aacgcctgaa atcccaagg caacagtgc 840

```

```

cgtcaggcat tcctttgtga aatcatcgac gcaggtaaga cacttgatcc tgcgaccggt 900
ggaaagtgcg tccatgacga aatccatcga ccagggtcaga ttgggcgcgc ccggacggag 960
cagcggcaga cgttctgttg ccagcccttt acgacgtctt ctgcgtttta cgcccaggcc 1020
actgagggtga taaagccggt acacgcgctt atgattaaca tgaagccctt cagggcgag 1080
caactgccaa atacgacggt agccaaaacg cctgcgctcc agtgccagct cagtgatgag 1140
ccctgataaa tgcgcatcag cagccggacg gtgagcctca tagcggcagg tgcacaggga 1200
taaacctgta agcctgcagg cacgacgttg cgacagaccg gtcgcatcac acatcaacat 1260
cacggcttcc cgcttctggt ctgtcgtcag tactttcgcc caagagccac ctgaagcgcc 1320
tcctttatcca gcatggcttc ggcaagcagc ttcttgagtc tgggtgttctc ttctcaagc 1380
gacttcaggc gcttaacttc aggcacctcc ataccgccat acttcttacg ccagggtgtaa 1440
aacgtggcat cggaaatggc atgcttgagg cagagttcac gggcggttac cccagcttcg 1500
gcttcgcgga gaatactgat gatctgttcg tcggaaaaac gcttcttcat ggggatgtcc 1560
tcatgtggct tatgaagaca ttactaacat cgggggtgtac taatcaacgg ggagcaggtc 1620
accatcacta tcaatctggc aaaaattgat gcgaaaactc tcggcctgga cggttttaat 1680
atcgatggcg cgcagaaagc aaccggcagt gacctgattt ctaaatttaa agcgacagg 1740
actgataatt atcaaatata cgttactgat aactatactg ttaatgtaga tagtggagta 1800
gtacaggata aagatggcaa acaagtttat gtgagtgtcg cggatgggtc acttacgacc 1860
agcagtgata ctcaattcaa gattgatgca actaagcttg cagtggctgc taaagattta 1920
gctcaaggta ataagattgt ctacgaagg atcgaattta caaataccgg cactggcgct 1980
atacctgccg caggtaatgg taaattaacc gccaatgttg atggtaaggc tgttgaattc 2040
actatttcgg ggagtgtgta tacatcaggg actagtgcga ccggtgcccc tacgacagcc 2100
ctatacaaaa atagtgcagg gcaattgact gcaacaaaag ttgaaaataa agcagcgaca 2160
ctatctgata ttgatctgaa cgctgccaa acgataacgg agactgcttc tggttaacat 2220
gcaacttacg atgttagtgc agatggtaaa ggtagcccca ttctggtaaa cgaagatgca 2340
aaagtcatgt atctgagcaa atcagaaggc ctcgaaaact tcgacaaaagc attggctaaa 2400
gcaaaaatcg tgcaatctac caccaccccg gtacaaaacc gtttcgactc tgccatcacc 2460
ggtgacaatc tgcgttctga cctcgggtga gtacaaaacc gtttcgactc tgccatcacc 2460
aaccttggca acaccgtaaa caacctgtct tctgcccgtg gccgtatcga agatgctgac 2520
tacgcgaccg aagtgtctaa catgtctcgt gcgcagatcc tgcaacaagc ggggtacctc 2580
gttctggcac aggctaacc 2599

```

```

<210> 53
<211> 1245
<212> DNA
<213> Escherichia coli

```

```

<400> 53
aacaacaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcccatt 60
aacagcgcta aagatgacgc tgcggggccag gcgattgcta accgcttcac ttctaacatc 120
aaaggtctga ctcaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180
gaaggcgcac tgtctgaaat caacaacaac ttgcagcggt ttcgtgaact gaccgttcag 240
gccactaccg gtactaactc tgattctgac ctgtcttcaa tccaggacga aatcaaatcc 300
cgtctcgatg aaattgaccg cgtatccggt cagactcagt tcaacggcgt gaacgtactg 360
gcaaaagatg gctcgatgaa aattcaggtc ggtgcaaatg atggtcagac aatcagcatt 420
gatttgacga agattgattc ttctacttta ggggttaaatg gtttttctgt ttcaaaaaat 480
gcagtatctg ttggtgatgc tattactcaa ttgcctggcg agacggcagc cgatgcacca 540
gtaaccatca agtttgatga ttcagtaaaa actgatttaa aactgaccga tgcttcaggg 600
ttaagtctgc ataacctcaa agatgaaaat ggtaatttaa ctaaccagta tgttgatcac 660
aatggcggaa aatcttacgc tgcacagtc gctgccaatg gtaatgttac gctgaacaaa 720
gcaaatgtaa cctacagcga tgctgcaaac ggtattgata ccgcaacgca gtcaggccag 780
ttagttcagg ttggtgcaga ttctaccggt acgcaaaaag cattcgtgtc tgtccaagg 840
aaaagctttg gcattgatga cgccgccttg aagaataaca ctggtgatgc taccgctact 900
ccaccgggaa catctgggac aacagttgtc gcagcgtcaa ttcactctgag tacgggcaaa 960
aactctgtag acgctgatgt aacggcttcc actgaattca cagggtgcttc aaccaacgat 1020
ccactgactc tgctggacaa agctatcgca tctgttgata aattccgttc ttctttgggg 1080
gcggtacaga accgtctgag ctccgctgta accaacctga acaacaccac caccacactg 1140
tctgaagcgc agtcccgtat tcaggacgcc gactatcgca ccgaagtgtc caacatgtcg 1200
aaagcgacga ttatccagca ggcaggtaac tccgtgctgt ccaaa 1245

```

<210> 54
 <211> 1212
 <212> DNA
 <213> Escherichia coli

<400> 54
 aacaaaaacc agtctgcgct gtcgacttct atcgaacgcc tctcttctgg cctgcgtatt 60
 aacagtgcga aagatgacgc tgccgggtcag gcgatagcta accgtttcac ctctaacatt 120
 aaaggcctga ctacaggctgc gcgtaacgcc aacgacggta tttctctggc gcagaccaca 180
 gaaggtgcgt tgtctgaaat caacaacaac ttgcaacgtg tgcgtgagtt gaccgttcag 240
 gcgacgaccg gtactaactc tgattctgac ctgtcatcta ttcaggacga aatcaaatcc 300
 cgtctggatg agattgaccg tgtttccggg cagacccagt tcaacggcgt gaatgtactg 360
 gcaaaagacg gttcgtatgaa gattcagggt ggcgcgaaatg atggccagac tattagcatt 420
 gatttacaga aaattgactc ttctacatta gggttgaatg gtttctccgt ttctgctcaa 480
 tcacttaacg ttggtgattc aattactcaa attacaggag ccgctgggac aaaacctgtt 540
 ggtgttgatt tcactgctgt tgcgaaagat ctgactactg cgacaggtaa aactgtcgat 600
 gtttccagcc tgacgttaca caacaccctg gatgcgaaag gggctgccac cgcacagttc 660
 gtcgttcaat ccggtagtga tttctactcc gcgtccattg accatgcaag tgggtgaagt 720
 acgttgataa aagccgatgt cgaatacaaa gacaccgata atggactaac gactgcagct 780
 actcagaaaag atcagctgat taaagttgcc gctgactctg acggcgcgcc tgccgggatat 840
 gtaacattcc agggtaaaaa ctacgctaca acggctccag cggcgcttaa tgatgacact 900
 acggcaacag ccacagcgaa caaagttggt gttgaattat ctacagcaac tccgactgcg 960
 cagttctcag gggcttcttc tgctgatcca ctggcacttt tagacaaagc cattgcacag 1020
 gttgatactt tccgctcctc cctcggtgcc gttcaaaacc gtctggactc tgcggtaacc 1080
 aacctgaaca acaccaccac caacctgtct gaagcgcagt cccgtattca ggacgcgcag 1140
 tatgcgaccg aagtgtctaa catgtcgaaa gcgcagatca tccagcaggc gggtaactct 1200
 tgctgtctta aa 1212

<210> 55
 <211> 1758
 <212> DNA
 <213> Escherichia coli

<400> 55
 atggcacaaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
 gcgaaggatg acgcgcggg tcaggcgatt gctaaccgtt ttactttctaa cattaaaggc 180
 ctgactcagg ctgcacgtaa cgccaacgac ggtattttctg ttgcacagac cactgaaggc 240
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggg tcaggcttct 300
 accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgctc 360
 gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
 gacggttcga tgaaaattca ggttggtgcg aatgacggtg aaactatcac tatcgacctg 480
 aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaaatgg taaaggtaact 540
 attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacacc 600
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
 gataaattag ggaatggcga taaagtcacc gttggcggcg tagattatac ttacaacgct 720
 aaatctgggtg attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780
 caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
 ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
 tcagcaggta atatcaccat cgggtggaagc caggcatacg tagacgatgc aggcaacttg 960
 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct taaagccgcg 1020
 agcgaaggta gtgacgggac ctctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
 gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
 gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
 attaccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
 gatgctgcga agtcttatgt ggatgataaa ggtggtatta ctaacgttgc cgactataca 1320
 gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
 actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440
 ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgcgtgc 1500
 ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
 cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgag 1620
 tccgttatcc aggcgcgcga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
 atccagcagg ccggttaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
 tctctgctgc agggtaa 1758

<210> 56
 <211> 14024
 <212> DNA
 <213> Escherichia coli

```

<400> 56
gtaaccaagg gcggtacgtg cataaatttt aatgcttattc aaaactatta gcattaaaaa 60
tatataagaa attctcaaatt gaacaaagaa accggtttcaa taattatgcc cgttttacaat 120
ggggccaaaa ctataatctc atcagtagaa tcaattatac atcaatctta tcaagattttt 180
gttttgtata tcattgacga ttgtagcacc gatgatacat tttcattaat caacagtcga 240
tacaaaaaca atcagaaaaat aagaatattg cgtaacaaga caaatttagg tgttgcagaa 300
agtcgaaatt atggaataga aatggccacg gggaaatata tttctttttg tgatgaggat 360
gatttgtggc acgagaaaaa attagagcgt caaatcgaag tgttaaataa tgaatgtgta 420
gatgtggtat gttctaatta ttatgttata gataacaata gaaatattgt tggcgaagtt 480
aatgctcctc atgtgataaa ttatagaaaa atgctcatga aaaactacat agggaaatttg 540
acaggaaatct ataattgcaa caaattgggt aagttttatc aaaaaaagat tggtcacgag 600
gattatttga tgtggctgga aataattaat aaaacaaatg gtgctatttg tattcaagat 660
aatctggcgt attacatgcg ttcaataat tctactatcg gtaataaaat taaagctgca 720
aaatggacat ggagtatata tagagaacat ttacatttgt cctttccaaa aacattatat 780
tattttttat tatatgcttc aaatggagtc atgaaaaaaa taacacattc actattaagg 840
agaaaggaga ctaaaaagtg aagtcagcgg ctaagttgat ttttttattc ctatttacac 900
tttatagtct ccagttgtat ggggttatca tagatgatcg tataacaaat tttgatacaa 960
aggtattaac tagtattata attatatttc agattttttt tgttttatta ttttatctaa 1020
cgattataaa tgaaagaaaa cagcagaaaa aattttatcgt gaactgggag cttaaagttaa 1080
tactcgtttt cctttttgtg actatagaaa ttgctgctgt agttttattt cttaaagaag 1140
gtattcctat atttgatgat gatccagggg gggctaaact tagaatagct gaaggtaactg 1200
gactttacat tagatattt aagtttttg gtaatatagt tgtgtttgca ttaattattc 1260
tttatgatga gcataaattc aaacagagga ccatcatatt tgtatatttt acaacgattg 1320
ctttatttgg ttatcgttct gaattgggtg tgctcattct tcaatatata ttgattacca 1380
atatacctgtc aaaggataac cgtaatccta aaataaaaaa aataataggg tattttttat 1440
tggtaggggt tgtatgctcg ttgttttattc taagtttagg acaagacgga gaacaaaatg 1500
actcatataa taatatgtta aggataatta ataggttaac aatagagcaa gttgaagggtg 1560
ttccatatgt tgtttctgaa tctattaaga acgatttctt tccgacacca gagttagaaa 1620
aggaattaaa agcaataata aatagaatac agggaaataa gcatcaagac ttattttatg 1680
gagaacggtt acataaacia gtatttggag acatgggagc aaatttttta tcagttacta 1740
cgtatggagc agaactgtta gttttttttg ttttctctg tgtattcatt atcccttttag 1800
ggatatatat acccttttat ctttttaaga gaatgaaaaa aacctatagc tcgataaatt 1860
gcgcattcta ttcatatattc attatgattt tattgcaata cttagtggct gggaatgcat 1920
cggccttctt ttttggctct tttctctcgg tattgataat gtgtactcct ctgatcttat 1980
tgcattgatac gttaaagaga ttatcacgaa atgaaaatat cagttataac tgtgacttat 2040
aataatgctt aagggttaga aaaaacttta agtagtttat caattttaaa aataaaacct 2100
tttgagatta ttatagttga tggcggctct acagatggaa cgaatcgtgt cattagtaga 2160
tttactagta tgaatattac acatgtttat gaaaaagatg aagggatata tgatgcatg 2220
aataagggcc gaatgttggc caaaggcgac ttaatacatt atttaaagc cggcgatagc 2280
gtaattggag atatatataa aaatatcaaa gagccatgtt tgattaaagt tggccttttc 2340
gaaaatgata aacttctggg attttctctc ataaccatt caaatacagg gtattgtcat 2400
caaggggtga ttttcccaaa gaatcattca gaatatgatc taaggatata aatatgtgct 2460
gattataagc ttattcaaga ggtgtttcct gaagggttaa gatctctatc tttgattact 2520
tcgggttatg taaaatatga tatgggggga gtatcttcaa aaaaaagaat ttttaagagat 2580
aaagagcttg ccaaaatttat gtttgaaaaa aataaaaaaa accttattaa gtttattcca 2640
atttcaataa tcaaaatttt attccctgaa cgtttaagaa gagtattgcy gaaaatgcaa 2700
tatatttgtc taactttatt cttcatgaag aatagttcac catatgataa tgaataaaat 2760
caaaaaaata cttaaatttt gcacttttaa aaaatatgat acatcaagtg ctttaggtag 2820
agaacaggaa aggtacagga ttatatcctt gtctgttatt tcaagtttga ttagtataat 2880
actctcacta ctttctctta tattaactgt aagtttaact ttaccttatt taggacaaga 2940
gagatttggg gtatggatga ctattaccag tcttgggtgct gctctgacat ttttggactt 3000
aggtatagga aatgcattaa caaacaggat cgcacattca tttgcgtgtg gcaaaaaattt 3060
aaagatgagt cggcaaatta gtgggtgggt cactttgctg gctggattat cgtttgtcat 3120
aactgcaata tgctatatta cttctggcat gattgatttg caactagtaa taaaagggtat 3180
aaacgagaat gtgtatgcag agttacaaca ctcaattaaa gtctttgtaa tcatatttgg 3240
acttggaaatt tattcaaatg gtgtgcacaa agtttatatg ggaatacaaa aagcctatat 3300
aagtaatatt gttaatgcc a ttttatattt gttatctatt attactctag taatatcgct 3360
gaaactacat gcgggactac cagttttaat tgtcagcact cttgggtattc aatacatatc 3420
gggaatctat ttaacaatta atcttattat aaagcgatta ataaagttta caaaaggttaa 3480
catacatgct aaaagagaag ctccatattt gatattaaac ggttttttct tttttatttt 3540
acagtttaggc actctggcaa catggagtgg tgataacttt ataatatcta taacattggg 3600
tgttacttat gttgctgttt ttagcattac acagagatta tttcaaatat ctacggtccc 3660

```

tcttacgatt	tataacatcc	cgttatgggc	tgettattgca	gatgctcatg	cacgcaatga	3720
tactcaat	ataaaaaaga	cgctcagaac	atcattgaaa	atagtgggta	tttcatcatt	3780
cttattggcc	ttcatattag	tagtggttcg	tagtggaagtc	gttaatat	ggacagaagg	3840
aaagattcag	gtacctcgaa	cattcataat	agcttatgct	ttatggctctg	ttattgatgc	3900
tttttcgaat	acatttgcaa	gctttttaaa	tggtttgaac	atagttaaac	aacaaatgct	3960
tgctgttgta	acattgatat	tgatcgcaat	tcagcaaaaa	tacatcatag	ttagccattt	4020
tgggttaact	gttatgttgt	actgcttcat	ttttatatat	attgtaaatt	acttttatatg	4080
gtataaatgt	agttttaaaa	aacatatcga	tagacagtta	aatataagag	gatgaaaatg	4140
aaatatatac	cagtttacca	accgtcattg	acaggaaaaag	aaaaagaata	tgtaaatgaa	4200
tgtctggagt	caacgtggat	ttcatcaaaa	ggaaactata	ttcagaagtt	tgaaaaataa	4260
tttgcggaac	aaaaccatgt	gcaatatgca	actactgtaa	gtaatggaa	ggttgctctt	4320
catttagctt	tgtagcggt	aggtatatcg	gaaggagatg	aagttattgt	tccaacactg	4380
acatatatag	catcagttaa	tgctataaaa	tacacaggag	ccacccccat	tttcgttgat	4440
tcagataatg	aaacttggca	aatgtctgtt	agtgcacatg	aacaaaaaat	cactaataaa	4500
actaaagcta	ttatgtgtgt	ccatttatat	ggacatccat	gtgatatgga	acaaattgta	4560
gaactggcca	aaagtagaaa	tttggttgta	attgaagatt	gcgctgaagc	ctttggttct	4620
aaatataaag	gtaaatatgt	gggaacattt	ggagatattt	ctacttttag	cttttttgga	4680
aataaaacta	ttactacagg	tgaagggtgga	atgggtgtca	cgaatgacaa	aacactttat	4740
gaccgttggt	tacatttttaa	aggccaagga	ttagctgtac	ataggcaata	ttggcatgac	4800
gttataggct	acaattatag	gatgacaaat	actgcgctg	ctataggatt	agcccagtta	4860
gaacaagctg	atgattttat	atcacgaaaa	cgtgaaattg	ctgatattta	taaaaaaat	4920
atcaacagtc	ttgtacaagt	ccacaaggaa	agtaaagatg	tttttcacac	ttattggatg	4980
gtctcaat	taactaggac	cgcagaggaa	agagagggaat	taagggaatca	ccttgccagat	5040
aaactcatcg	aaacaaggcc	agttttttac	cctgtccaca	cgatgccaat	gtactcggaa	5100
aaatatcaaa	agcaccctat	agctgaggat	cttgggtggc	gtggaattaa	tttacctagt	5160
ttccccagcc	tatcgaaatga	gcaagtttatt	tatatattgtg	aatctatttaa	cgaattttat	5220
agtataaat	agcctaaaat	attgtaaagg	tcattcatga	aaattgcgtt	gaattcagat	5280
ggattttacg	agtggggcgg	tgggaattgat	tttattaaat	atattctgtc	aatattagaa	5340
acgaaaccag	aaatatgtat	cgatattcct	ttaccgagaa	atgatataca	ttctcttata	5400
agagaaaaag	catttccttt	taaaagtata	ttaaaagcaa	ttttaaagag	ggaaaggcct	5460
cgatggattt	cattaaatag	atttaattgag	caatactata	gagatgcctt	tacacaaaa	5520
aatatagaga	cgaatcttac	ctttattaaa	agtaagagct	ctgcctttta	ttcatatttt	5580
gatagtagcg	attgtgatgt	tattcttctt	tgcatgcgtg	ttccttcggg	aaatttgaat	5640
aaaaaagcat	ggattgggtta	tatttatgac	tttcaacact	gttactatcc	ttcatttttt	5700
agtaagcgag	aaatagatca	aaggaatgtg	ttttttaaat	tgatgctcaa	ttgcgcatac	5760
aatattattg	ttaatgcaca	ttcagttatt	accgatgcaa	ataaatatgt	tgggaattat	5820
tctgcaaaaac	tacattctct	tccatttagt	ccatgccctc	aattaaaaatg	gttcgctgat	5880
tactctggta	atattgccaa	atataatatt	gacaaggatt	atttttataat	ttgcaatcaa	5940
ttttggaaac	ataaagatca	tgcaactgct	tttagggcat	ttaaaattta	tactgaatat	6000
aatcctgatg	tttatttagt	atgcacggga	gctactcaag	attatcgatt	ccctggatat	6060
tttaatgaat	tgatggtttt	ggcaaaaaag	ctcggaattg	aatcgaaaa	taagattatta	6120
gggcatat	ctaaacttga	acaaattgaa	ttaatcaaaa	attgcattgc	tgtaatacaa	6180
ccaaccttat	ttgaaggcgg	gcctggagg	ggggtaacat	ttgacgctat	tgcattaggg	6240
aaaaaagtta	tactatctga	catagatgtc	aataaaagaag	ttaatgtcgg	tgatgtatat	6300
ttcttttcagg	caaaaaacca	ttattcatta	aatgacgcga	tggtaaaagc	tgatgaatct	6360
aaaatttttt	atgaacctac	aactctgata	gaattgggtc	tcaaaaagacg	caatgcgtgt	6420
gcagattttt	tttttagatgt	tgtgaaacaa	gaaattgaat	cccgatctta	atatattcaa	6480
gaggtatata	atgactaaag	tcgctcttat	tacaggtgta	actggacaag	atggatctta	6540
tctagctgag	tttttgcttg	ataaagggtta	tgaagttcat	ggtatcaaac	gccgagcctc	6600
atctttttaat	acagaacgca	tagaccatat	ttatcaagat	ccacatgggt	ctaaccctaa	6660
ttttcacttg	cactatggag	atctgactga	ttcatctaac	ctcactagaa	ttctaaagga	6720
ggtacagcca	gatgaagtat	ataatttagc	tgctatgagt	cacgtagcag	tttcttttga	6780
gtctccagaa	tatacagccg	atgtcgatgc	aattggtaca	ttacgtttac	tgggaagcaat	6840
tcgcttttta	ggattggaaa	acaaaaacg	tttctatcaa	gcttcaacct	cagaattata	6900
tggacttggt	caggaaatcc	ctcaaaaaga	atccaccttc	ttttatcttc	gttcccctta	6960
tgcagttgca	aaactttacg	catattggat	accggtaaat	tatcgagagt	catatgggtat	7020
ttatgcatgt	aatgggtatat	tgttcaatca	tgaatctcca	cgccgtggag	aaacgtttgt	7080
aacaaggaaa	attactcgag	gacttgcaaa	tattgcacaa	ggcttggaat	catgtttgta	7140
tttagggaat	atggattcgt	tacgagattg	gggacatgca	aaagattatg	ttagaatgca	7200
atggttgatg	ttacaacagg	agcaaccgca	agattttgtg	attgcaacag	gagtccaata	7260
ctcagtcggt	cagtttgcg	aaatggcagc	agcacaaact	ggtattaaga	tgagctttgt	7320
tggtaaaagga	atcgaagaaa	aaggcattgt	agattcggtt	gaaggacagg	atgctccagg	7380
tgtgaaacca	ggtgatgtca	ttgttgctgt	tgatcctcgt	tatttccgac	cagctgaagt	7440
tgatactttg	cttgagatgc	cgagcaaaagc	taatctcaaa	cttgggttga	gaccagaaat	7500
tactcttgct	gaaatgattt	ctgaaatggt	tgccaaagat	cttgaagccg	ctaaaaaaca	7560
ttctctttta	aaatcgcatg	gtttttctgt	aagcttagct	ctggaatgat	gatgaataag	7620
caacgtattt	ttattgctgg	tcaccaagga	atgggtggat	cagctattac	ccgacgcctc	7680

aaacaacgtg	atgatgttga	gttggtttta	cgtactcggg	atgaattgaa	cttggtggat	7740
agtagcgtg	ttttggattt	tttttcttca	cagaaaatcg	accagggtta	tttggcagca	7800
gcaaaagtcg	gaggtatttt	agctaacagt	tcttatcctg	ccgattttat	atatgagaat	7860
ataatgatag	aggcgaaatgt	cattcatgct	gcccacaaaa	ataatgtaaa	taaactgctt	7920
ttcctcgggt	cgtcgtgtat	ttatcctaag	ttagcacacc	aaccgattat	ggaagacgaa	7980
ttattacaag	ggaaacttga	gccaacaaat	gaaccttatg	ctatcgcaaa	aattgcaggt	8040
attaaattat	gtgaatctta	taaccgtcag	tttgggcgtg	attaccgttc	agtaatgcca	8100
accaatcttt	atggtcctaaa	tgacaatttt	catccaagta	attctcatgt	gattccggcg	8160
cttttgcgcc	gctttcatga	tgctgtggaa	aacaattctc	cgaatgttgt	tgtttgggga	8220
agtggactc	caaagcgtga	attcttacat	gtagatgata	tggcttctgc	aagcatttat	8280
gtcatggaga	tgccatacga	tatatggcaa	aaaaatacta	aagtaatgtt	gtctcatatc	8340
aatattggaa	caggtattga	ctgcacgatt	tgtgagcttg	cggaaacaat	agcaaaagtt	8400
gtaggttata	aagggcatat	tacgttcgat	acaacaaagc	ccgatggagc	ccctcgaaaa	8460
ctacttgatg	taacgcttct	tcatcaacta	ggttggaatc	ataaaaattac	ccttcacaag	8520
ggtcttgaaa	atacatataa	ctggtttctt	gaaaaccaac	ttcaatatcg	ggggtaataa	8580
tggtttttaca	ttcccaagac	tttgccacaa	ttgtaaggtc	tactcctctt	atttctatag	8640
atttgattgt	ggaaaaacgag	tttggcgaaa	ttttgctagg	aaaacgaatc	aaccgcccgg	8700
cacagggcta	ttggttcgtt	cctgggtggta	gggtgttgaa	agatgaaaaa	tgccagacag	8760
cctttgaacg	attgacagaa	attgaaactag	gaattcgttt	gcctctctct	gtgggtaagt	8820
tttatgggat	ctggcagcac	ttctacgaag	acaatagtat	ggggggagac	ttttcaacgc	8880
attatatagt	tatagcattc	cttcttaaat	tacaaccaa	cattttgaaa	ttaccgaagt	8940
cacaacataa	tgcttattgc	tggtctatcg	gagcaaagct	gataaatgat	gacgatgtgc	9000
attataattg	tcgcgcata	tttaacaata	aaacaaatga	tgcgattggc	ttagataata	9060
aggatataat	atgtctgatg	cgccaataat	tgctgtagtt	atggccgggtg	gtacaggcag	9120
tcgtcttttg	ccactttctc	gtgaactata	tccaaagcag	tttttacaac	tctctgggtga	9180
taacaccttg	ttacaaaacga	cttttgctacg	acttttcaggc	ctatcatgtc	aaaaaccatt	9240
agtgataaca	aatgaacagc	atcgctttgt	tgtggctgaa	cagttaaggg	aaataaataa	9300
attaaatggt	aatattattc	tagaaccatg	cgggcgaaat	actgcaccag	caatagcgat	9360
atctcgcttt	catgcgttaa	aacgtaatcc	tcagggaagat	ccattgcttc	tagttctttg	9420
ggcagaccac	gttatagcta	aagaaaagtgt	tttctgtgat	gctattaaaa	atgcaactcc	9480
catcgcta	caaggtaaaa	ttgtaacggt	tggaattata	ccagaatatg	ctgaaactgg	9540
ttatgggtat	attgagagag	gtgaactatc	tgtaccgctt	caagggcgatg	aaaatactgg	9600
tttttattat	gtaaataagt	ttgtcgaaaa	gcctaactcg	gaaaccgcag	aattgtatat	9660
gacttctggt	aatcactatt	ggaatagtgg	aatattcatg	tttaaggcat	ctgtttatct	9720
tgaggaattg	agaaaaattta	gacctgacat	ttacaatgtt	tgtgaacagg	ttgcctcatc	9780
ctcatacatt	gatctagatt	ttattcgatt	atcaaaagaa	caatttcaag	attgtcctgc	9840
tgaatctatt	gattttgctg	taatggaaaa	aacagaaaaa	tgtgttgtat	gccctgttga	9900
tattgggttg	agtgcgcttg	gatcttgcca	atcggttatg	gacatttagtc	taaaatcgaa	9960
aacaggagat	gtatgtaaag	gtgatataat	aacctatgat	actaagaata	attatatcta	10020
ctctgagtc	gcgttggtag	ccgccattgg	aattgaagat	atggttatcg	tgcaactcaa	10080
agatgcccgt	cttgtgtcta	aaaagagtga	tgtacagcat	gtaaaaaaa	tagtcgaaat	10140
gcttaaatgt	cagcaacgta	cagagtatat	tagtcatcgt	gaagttttcc	gaccatgggg	10200
aaaatttgat	tcgattgacc	aaggtgagcg	atacaaaagtc	aagaaaatta	ttgtgaaacc	10260
tggtgagggg	ctttcttttaa	ggatgcatca	ccatcgcttc	gaacatttga	tcgtgctttc	10320
tggtcacagca	aaagtaaccc	ttggcgataa	aactaaacta	gtcaccgcaa	atgaatcgat	10380
atacattccc	cttggcgag	cgtatagtct	tgagaatccg	ggcataatcc	ctcttaattct	10440
tattgaagtc	agttcagggg	attatttggg	agaggatgat	attataagac	agaaagaacg	10500
ttacaacat	gaagattaac	atatgaaatc	tttaacctgc	tttaaaacct	atgatattcg	10560
cgggaaatta	ggcgaagaac	tgaatgaaga	tattgcctgg	cgcattgggc	gtgcctatgg	10620
cgaattttctc	aaaccgaaaa	ccattgtttt	aggcgggtgat	gtccgcctca	ccagcgaagc	10680
gttaaaaactg	gcgcttgcca	aagggtttaca	ggatgcgggc	gtcgatgtgc	tggatatcgg	10740
tatgtccggc	accgaagaga	tctatttctc	cacgttccat	ctcggagtgg	atggcggcat	10800
cgaagtacc	gccagccata	accgatggga	ttacaacggc	atgaagctgg	tgccggaagg	10860
ggctcgcccg	atcagcgggtg	ataccggact	gcgcgatgtc	cagcgtctgg	cagaagccaa	10920
tgacttccct	cctgtcgatg	aaaccaaacc	tggctcgctat	cagcaaatca	atctgcgtga	10980
cgcttacggt	gatcacctgt	tcggttatat	caacgtcaaa	aacctcacgc	cgctcaagct	11040
ggtgatcaac	tccgggaacg	gcgcagcggg	tccgggtggtg	gacgccattg	aagcccgaatt	11100
taaagccctc	ggcgacccgg	tggaaattaat	caaagtacac	aacacgccgg	acggcaattt	11160
ccccaacggg	attcctaacc	cgtcgtctgc	ggaatgccgc	gacgacaccc	gtaatgcggg	11220
catcaaacac	ggcgcgggata	tgggcattgc	ctttgatggc	gattttgacc	gctgtttcct	11280
gtttgacgaa	aaagggcagt	ttatcgaggg	ctactacatt	gtcggcctgc	tggcagaagc	11340
gttctctgaa	aaaaatcccc	gcgcgaagat	catccacgat	ccacgtctct	cctggaacac	11400
cgttgatgtg	gtgactgccg	caggcggcac	cccggtaatg	tcgaaaaccg	gacacgcctt	11460
tattaaagaa	cgtatgcgca	aggaagacgc	catctacggt	ggcgaaatga	gcgctcacca	11520
ttacttccgt	gatttctgctt	actgcgcacg	cggcatgac	ccgtggctgc	tggctgcgca	11580
actgggtgtg	ctgaaaggaa	aaacgctggg	cgcgaaccgga	tggcggcggt	11640	
tccggcaagc	ggtgagatca	acagcaaaact	ggcgcaacc	gttgaggcaa	ttaatcgctg	11700

ggaacagcat	tttagccg	aggcgctggc	ggtggatcgc	accgatggca	tcagcatgac	11760
ctttgccc	tgccgcttta	acctgcgctc	ctccaacacc	gaaccggtgg	tgccggttgaa	11820
tgtggaatca	cgcggtgatg	taaaagcta	ggaaaagaaa	actaaagctc	ttcttaaatt	11880
gctaagttag	tgattattta	cattaatcat	taagcgtatt	taagattata	ttaaagta	11940
gttattgcgg	tatatgatga	atatgtgggc	ttttttatgt	ataacgacta	taccgcaact	12000
ttatctagga	aaagattaat	agaaataaag	ttttgtactg	accaatttgc	atttcacgct	12060
acgattgaga	cgttcctttg	cttaagacat	tttttcatcg	cttatgta	aacaaatgtg	12120
ccttatataa	aaaggagaac	aaaaatggaa	ttaaaataat	tgagacaata	gattttttatt	12180
atccctgttt	acgatattat	agccaaagtt	gtatcctgca	tcagtcctgc	aatatttcac	12240
gagtgccttg	ttaactgaat	acatgtctgc	cattttccag	atgataacga	cgtcacgcga	12300
attgatggta	aaacacttcg	gcacacttat	gacaagagtc	gtcgcagagg	agtgggttcat	12360
gtcatttagt	cgtttcagca	atgcacagtc	tggtcctcgg	atagatcaag	acggatgaga	12420
aacctaatac	gttcacagtt	attcatgaac	tttctaaaat	gatgggtatt	aaaggaaaaa	12480
taatacatac	tgatgcgatg	gcttgccaga	aagatattgc	agagaagata	taaaaacaga	12540
gatgtgatta	tttattcgct	gtaaaaggaa	ataagagtcg	gcttaataga	gtctttgagg	12600
agatattttac	gctgaaagaa	ttaataatc	caaaacatga	cagttacgca	attagtga	12660
agaggcacgg	cagagacgat	gtccgtcttc	atattgtttg	agatgctcct	gatgagctta	12720
ttgatttcac	gtttgaatgg	aaagggtctg	agaattttat	aatggcagtc	cactttctct	12780
caataatagc	agagcaaaag	aaagaatccg	aatgacgat	caaataattat	attagatctg	12840
ctgctttaac	cgcagagaag	ttcgccacag	taaatcgaaa	tcactggcgc	atggagaata	12900
agttgcacac	tagcctgatg	tggtaatgaa	tgaaatcgac	tataatataa	gaaggcgagt	12960
tgcatctgaa	tgattttcta	gaatgcggca	catcgctatt	aatatctgac	aatgataatg	13020
tattcaaggc	aggattatca	tgtaagatgc	gaaaagcagt	catggacaga	aacttcctag	13080
cgtcaggcat	tgacgcgtgc	gggctttcat	aatcttgc	tggttttgat	aagatatttc	13140
tttgagagat	ggaaaatgaa	tttgtatgg	atttttgg	ctggaagtta	tggtagagaa	13200
acaataacca	ttctaaatca	acaaataaag	caagaatgtg	gttctgacta	tgctctggtt	13260
tttgtggatg	atgttttggc	aggaaagaaa	gttaatgg	ttgaagtgt	ttcaaccaac	13320
tgctttctaa	aagccctta	tttaaaaaag	tattttta	ttgctattgc	taatgataag	13380
atacgacaga	gagtgtctga	gtcaatatta	ttacacggg	ttgaaccaat	aactataaaa	13440
catccaaata	gcgttgttta	tgatcatact	atgataggta	gtggcgctat	tatttctccc	13500
tttggttaca	tatctactaa	tactcatata	gggaggtttt	ttcatgcata	catatactca	13560
tacgttgcac	atgattgtca	aataggagac	tatgttacat	ttgctcctgg	ggctaaatgt	13620
aatggatatg	ttgttattga	agacaatgca	tatataggct	cggtgacagt	aattaagcag	13680
ggtgttccta	atcgcccaact	tattattggc	gcgggagcca	ttatagggtat	gggggctgtt	13740
gtcactaaaa	gtgttcctgc	cggtataact	gtgtgcggaa	atccagcaag	agaaatgaaa	13800
agatcgccaa	catctattta	atgggaatgc	gaaaacacgt	tccaaatggg	actaatgttt	13860
aaaaatatata	taatttcgct	aatttactaa	attatggctt	ctttttaagc	tatcctttac	13920
ttagttatta	ctgatacagc	atgaaattta	taatactctg	atacattttt	atagcttatt	13980
caagccgcat	atctagcgg	aacccctgac	aggagtaaac	aatg		14024

<210> 57

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 57

atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagctct	cgctgtcgag	ttctatcgag	cgtctgtctt	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgccgcagg	tcaggcgatt	gctaaccgtt	ttacttctaa	cattaaaggc	180
ctgactcagg	cgcccgtaaa	cgccaacgac	ggtatttctg	ttgcgcagac	caccgaaggc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgtattcgtg	aactgacggg	tcaggccact	300
acagggacta	actccgattc	tgacctggac	tccatccagg	acgaaatcaa	atctcgtctt	360
gatgaaattg	accgcgtatc	cggccagacc	cagttcaacg	gcgtgaacgt	gctggcgaaa	420
gacggttcaa	tgaaaattca	ggttgggtgc	aatgacggcg	aaaccatcac	gatcgacctg	480
aaaaaaaaat	attctgatac	tctgggtctg	aatggcttta	acgtaaatgg	taaaagtact	540
attaccaaca	aagctgcaac	ggtaagtgat	ttaacttctg	ctggcgcgaa	gttaaacacc	600
acgacaggtc	tttatgatct	gaaaaccgaa	aataccttgt	taactaccga	tgctgcattc	660
gataaattag	ggaatggcga	taaagtcaca	gttggcgcg	tagattatac	ttacaacgct	720
aaatctgggt	attttactac	cactaaatct	actgctggta	cgggtgtaga	cgcgcggcgg	780
caggctgctg	attcagcttc	aaaacgtgat	gcgttagctg	ccacccttca	tgctgatgtg	840
ggtaaatctg	ttaatggttc	ttacaccaca	aaagatggta	ctgtttcttt	cgaacaggat	900
tcagcaggta	atatcaccat	cggtggaagc	caggcatacg	tagacgatgc	aggcaacttg	960
acgactaaca	acgctggtag	cgcagctaaa	gctgatatga	aagcgctgct	caaagcagcg	1020
agcgaaggta	gtgacgggtg	ctctctgaca	ttcaatggca	cagaatatac	catcgcaaaa	1080
gcaactcctg	cgacaaccac	tccagtagct	cgggttaatcc	ctgggtggg	tacttatcag	1140
gctacacgtg	gtaaagatgt	agtattgagc	gaaaccaaag	cggctgcgcg	gacatcttca	1200
attaccttta	attccggtgt	actgagcaaa	actattgggt	ttaccgcggg	tgaatccagt	1260

gatgctgcga	agtcttatgt	ggatgataaa	ggtggtatca	ctaacgttgc	cgactataca	1320
gtctcttaca	gcgttaacaa	ggataacggc	tctgtgactg	ttgccgggta	tgcttcagcg	1380
actgatacca	ataaagatta	tgctccagca	attggtactg	ctgtaaattgt	gaactccgcg	1440
ggtaaaatca	ctactgagac	taccagtgtc	ggttctgcaa	cgaccaaccc	gcttgctgcc	1500
ctggacgacg	caatcagctc	catcgacaaa	ttccggttct	ccctgggtgc	tatccagaac	1560
cgtctggatt	ccgcagtcac	caacctgaac	aacaccacta	ccaacctgtc	cgaagcgacg	1620
tcccgatttc	aggacgccga	ctatgcgacc	gaagtgtcca	acatgtcgaa	agcgcagatc	1680
attcagcagg	ccggtaactc	cgtgctggca	aaagctaacc	aggtaccgca	gcaggttctg	1740
tctctgctgc	agggttaa					1758

<210> 58
 <211> 1758
 <212> DNA
 <213> Escherichia coli

<400> 58						
atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagtctg	cgctgtcgag	ttctatcgag	cgtctgtctt	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgcagcggg	tcaggcgatt	gctaaccggt	ttacttctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaacgac	ggtatttctg	ttgcacagac	cactgaaggc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgtattcgtg	aactgacggg	tcaggccact	300
acagggacta	actccgattc	tgacctggac	tccatccagg	acgaaatcaa	atctcgtctt	360
gatgaaattg	accgcgtatc	cggccagacc	cagttcaacg	gcgtgaacgt	gctggcgaaa	420
gacggttcaa	tgaaaattca	ggttgggtgcg	aatgacggcg	aaaccatcac	gatcgacctg	480
aaaaaaatcg	attctgatac	tctgggtctg	aatggcttta	acgtaaatgg	taaagggtact	540
attaccaaca	aagctgcaac	ggtaagtgat	ttaacttctg	ctggcgcgaa	gttaaaccacc	600
acgacaggtc	tttatgatct	gaaaaccgaa	aataccttgt	taactaccga	tgctgcattc	660
gataaattag	ggaatggcga	taaagtcaca	gttggcggcg	tagattatac	ttacaacgct	720
aaatctgggt	attttactac	cactaaatct	actgctggta	cgggtgtaaa	cgccgcggcg	780
caggctgctg	attcagcttc	aaaacgtgat	gcgttagctg	ccacccttca	tgctgatgtg	840
ggtaaatctg	ttaatggttc	ttacaccaca	aaagatggta	ctgtttcttt	cgaaacggat	900
tcagcaggta	atatcaccat	cgggtggaagc	caggcatacg	tagacgatgc	aggcaacttg	960
acgactaaca	acgctggtag	cgcagctaaa	gctgatatga	aagcgtgct	caaagcagcg	1020
agcgaaggta	gtgacgggtc	ctctctgaca	ttcaatggca	cagaatatac	catcgcaaaa	1080
gcaactcctg	cgacaaccac	tccagtagct	ccgttaatcc	ctggtgggat	tacttatcag	1140
gttacagtga	gtaaagatgt	agtattgagc	gaaaccaaag	cggctgccgc	gacatcttca	1200
attaccttta	attccgggtg	actgagcaaa	actattgggt	ttaccgcggg	tgaatccagt	1260
gatgctgcga	agtcttatgt	ggatgataaa	ggtggtatca	ctaacgttgc	cgactataca	1320
gtctcttaca	gcgttaacaa	ggataacggc	tctgtgactg	ttgccgggta	tgcttcagcg	1380
actgatacca	ataaagatta	tgctccagca	attggcactg	ctgtaaattgt	gaactccgcg	1440
ggtaaaatca	ctactgagac	taccagtgtc	ggttctgcaa	cgaccaaccc	gcttgctgcc	1500
ctggacgacg	caatcagctc	catcgacaaa	ttccggttct	ccctgggtgc	tatccagaac	1560
cgtctggatt	ccgcggtcac	caacctgaac	aacaccacta	ccaacctgtc	cgaagcgacg	1620
tcccgatttc	aggacgccga	ctatgcgacc	gaagtgtcca	acatgtcgaa	agcgcagatc	1680
atccagcagg	ccggtaactc	cgtgctggca	aaagctaacc	aggtaccgca	gcaggttctg	1740
tctctgctgc	agggttaa					1758

<210> 59
 <211> 1758
 <212> DNA
 <213> Escherichia coli

<400> 59						
atggcacaag	tcattaatac	caacagcctc	tcgctgatca	ctcaaaataa	tatcaacaag	60
aaccagtctg	cgctgtcgag	ttctatcgag	cgtctgtctt	ctggcttgcg	tattaacagc	120
gcgaaggatg	acgcgcgggg	tcaggcgatt	gctaaccggt	ttacttctaa	cattaaaggc	180
ctgactcagg	ctgcacgtaa	cgccaacgac	ggtatttctg	ttgcacagac	cactgaaggc	240
gcgctgtccg	aaatcaacaa	caacttacag	cgtatccgtg	agctgacggg	tcaggcttct	300
accgggacta	actctgattc	ggatctggac	tccattcagg	acgaaatcaa	atcccgctct	360
gacgaaaattg	accgcgtatc	cggtcagacc	cagttcaacg	gcgtgaacgt	actggcaaaa	420
gacggttcga	tgaaaattca	ggttgggtgcg	aatgacgggtg	aaactatcac	tatcgacctg	480
aagaaaaatcg	attctgatac	tctgggtctg	aatgggttta	acgtaaatgg	taaagggtact	540
attaccaaca	aagctgcaac	ggtaagtgat	ttaacttctg	ctggcgcgaa	gttaaaccacc	600
acgacaggtc	tttatgatct	gaaaaccgaa	aataccttgt	taactaccga	tgctgcattc	660
gataaattag	ggaatggcga	taaagtcacc	gttggcggcg	tagattatac	ttacaacgct	720
aaatctgggtg	atcttactac	caccaaactct	actgctggta	cgggtgtaga	cgccgcggcg	780


```

caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaactctg ttaatgggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct taaagccgcg 1020
agcgaaggta gtgacggtgc ctctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtgggtatta ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtt ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
atccagcagg ccggttaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa 1758

```

<210> 60

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 60

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaagggc 180
ctgactcagg cggcccgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggg tcaggccact 300
acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360
gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacggttcaa tgaaaattca ggttgggtgc aatgacggcg aaaccatcac gatcgacctg 480
aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaatgg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacacc 600
acgacaggtc ttatgatctt gaaaaccgaa aataccttgt taactacgga tgctgcattc 660
gataaaattg ggaatggcga taaagtcaca gttggcggcg tagattatac ttacaacgct 720
aaatctgggtg attttactac cactaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctgctg attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatgggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgctt caaagcagcg 1020
agcgaaggta gtgacggtgc ctctctgaca ttcaatggca cagaatatac catcgcaaaa 1080
gcaactcctg cgacaaccac tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtgggtatc ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtt ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg caatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc cgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
attcagcagg ccggttaactc cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa 1758

```

<210> 61

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 61

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaagggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggg tcaggccact 300

```

```

acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360
gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacgggttcaa tgaaaattca ggttggtgcg aatgacggcg aaaccatcac gatcgacctg 480
aaaaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaatgg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacacc 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcaca gttggcggcg tagattatac ttacaacgct 720
aaatctgggtg attttactac cactaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctgctg attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct caaagcagcg 1020
agcgaaggta gtgacgggtg ctctctgaca ttcaatggca cagaatatac catcgcaaaa 1080
gcaactcctg cgacaaccac tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgca agtcttatgt ggatgataaa ggtggtatca ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggcactg ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaacc gcttgctgcc 1500
ctggacgacg caatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcgggtc caacctgaac aacaccacta ccaacctgtc cgaagcgag 1620
tcccgatttc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgagatc 1680
atccagcagg ccggtaaactc cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa

```

<210> 62

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 62

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcggg tcaggcgatt gctaaccggt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgctc 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaaaattca ggttggtgcg aatgacgggt aaactatcac tatcgacctg 480
aagaaaatcg attctgatac tctgggtctg aatggtttta acgtaaatgg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacacc 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcacc gttggcggcg tagattatac ttacaacgct 720
aaatctgggtg attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct taaagccgcg 1020
agcgaaggta gtgacgggtg ctctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctggtgggat ttcttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgca agtcttatgt ggatgataaa ggtggtatta ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtagct ctgtaaatgt gaactccgcg 1440
ggtaaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaacc gcttgctgcc 1500
ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgag 1620
tcccgatttc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgagatt 1680
atccagcagg ccggtaaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa

```

<210> 63
 <211> 1758
 <212> DNA
 <213> Escherichia coli

<400> 63
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
 gcgaaggatg acgccgcagg tcaggcgatt gctaaccggt ttactttctaa cattaagggc 180
 ctgactcagg cggcccgtaa cgccaacgac ggtatttctg ttgcccagac caccgaaggc 240
 gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggg tcaggccact 300
 acaggggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360
 gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
 gacggttcaa tgaaaattca ggttggtgcg aatgacggcg aaaccatcac gatcgacctg 480
 aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaatgg taaagggtact 540
 attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaaccac 600
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
 gataaattag ggaatggcga taaagtcaca gttggcggcg tagattatac ttacaacgct 720
 aaatctgggtg attttactac cactaaatct actgctggtg cgggtgtaga cgccgcggcg 780
 caggctgctg attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
 ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
 tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgctt caaagcagcg 1020
 agcgaaggta gtgacgggtg ctctctgaca ttcaatggca cagaatatac catcgcaaaa 1080
 gcaactcctg cgacaaccac tccagtagct ccgttaatcc ctgggtggat tacttatcag 1140
 gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgcccgc gacatcttca 1200
 attaccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
 gatgctgcga agtcttatgt ggatgataaa ggtgggtatc ctaacgttgc cgactataca 1320
 gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
 actgatacca ataaagatta tgctccagca attgggtact ctgtaaatgt gaactccgcg 1440
 ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
 ctggacgacg caatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
 cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc cgaagcgag 1620
 tcccgtattc aggcgcgca ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1680
 attcagcagg ccggtaaact cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740
 tctctgctgc aggtttaa 1758

<210> 64
 <211> 1758
 <212> DNA
 <213> Escherichia coli

<400> 64
 atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
 gcgaaggatg acgccgcggg tcaggcgatt gctaaccggt ttactttctaa cattaagggc 180
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcccagac caccgaaggc 240
 gcgctgtctg aaatcaacaa caacttacag cgtatccgtg agctgacggg tcaggcttct 300
 accggaacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctt 360
 gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt actggcaaaa 420
 gacggttcga tgaaaattca ggttggtgcg aatgacgggt aaactatcac tatcgacctg 480
 aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaaatgg taaagggtact 540
 attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaaccac 600
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
 gataaattag ggaatggcga taaagtcaac gttggcggcg tagattatac ttacaacgct 720
 aaatctgggtg attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780
 caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
 ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
 tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgctt taaagccgct 1020
 agcgaaggta gtgacgggtg ttctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
 gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctgggtggat tacttatcag 1140
 gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgcccgc gacatcttca 1200
 attaccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
 gatgctgcga agtcttatgt ggatgataaa ggtgggtatta ctaacgttgc cgactataca 1320
 gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380

```

actgatacca ataaagatta tgctccagca attggtactg ctgtaaattgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtt ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
atccagcagg ccggttaact cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa
1758

```

```

<210> 65
<211> 1758
<212> DNA
<213> Escherichia coli

```

```

<400> 65
atggcacaag tcattaatac caacagcctc tgcgtgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaaggcg 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggg tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaaaattca ggttggtgcg aatgacgggt aaactatcac tatcgacctg 480
aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaatgg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacacc 600
acgacagggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaaattg ggaatggcga taaagtcacc gttggcgcg tagattatac ttacaacgct 720
aaatctgggtg attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcacctc cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgctg taaagccgcg 1020
agcgaaggta gtgacgggtg ctctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctgggtggggt ttcttatcag 1140
gctacagtga gtaaaagtgt agtattgagc gaaaccaaag cggctgcccgc gacatcttca 1200
attaccttta attccgggtg actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatctgcgca agtcttatgt ggaatgataaa ggtgggtatta ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtactg ctgtaaattgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtt ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
atccagcagg ccggttaact cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa
1758

```

```

<210> 66
<211> 1788
<212> DNA
<213> Escherichia coli

```

```

<400> 66
atggcacaag tcattaatac caacagcctc tgcgtgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaaggcg 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggg tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaaaattca ggtaggtgcg aacgacggcc agactatcac tattgatctg 480
aagaaaattg actctgatac gctggggctg aatgggttta acgtgaatgg ttccggtagc 540
atagccaata aagcggcgac cattagcgac ctgacagcag cgaaaatgga tgctgcaact 600
aatactataa ctacaacaaa taatgcgctg actgcatcaa aggcccttga tcaactgaaa 660
gatggtgaca ctgttactat caaagcagat gcagctcaaa ctgccacggg ctatacatac 720
aatgcatctg ctggttaact ctcattcagt aatgtatcga ataatacttc agcaaaagca 780
ggtgatgtag cagctagcct tctcccgcgc gctgggcaaa ctgctagtgg tgtttacaaa 840
gcagcaagcg gtgaagtga ctttgatgtt gatgcgaatg gtaaaattac aatcggagga 900
caggaagcct atttaactag tgatggtaac ttaactacaa acgatgctgg tgggtgcgact 960

```

```

gcggtacgc ttgatggttt attcaagaaa gctggtgatg gtcaatcaat cgggtttaat 1020
aagactgcat cagtcacgat ggggggaaca acttataact ttaaaacggg tgctgatgct 1080
ggtgctgcaa ctgctaacgc aggggtatcg ttactgata cagctagcaa agaaaccgtt 1140
ttaaataaag tggctacagc taaacaaggc acagcagttg cagctaacgg tgatacatcc 1200
gcaacaatta cctataaatc tggcgttcag acgtatcagg cggtatattgc cgcaggtgac 1260
ggtactgcta gcgcaaaata tgccgataat actgacgttt ctaatgcaac agcaacatac 1320
acagatgctg atggtgaaat gactacaatt ggttcataca ccacgaagta ttcaatcgat 1380
gctaacaacg gcaaggtaac tgttgattct ggaactgggt cgggtaaata tgcgcgaaa 1440
gtcggggctg aagtatatgt tagtgctaatt ggtactttaa caacagatgc aactagcgaa 1500
ggcacagtaa caaaagatcc actgaaagct ctggatgaag ctatcagctc catcgacaaa 1560
ttccgttcat ccctgggggc tatccaaaac cgtttggatt ccgccgtcac caacctgaac 1620
aacaccacta ccaacctgtc tgaagcgcag tcccgtattc aggacgccga ctatgcgacc 1680
gaagtgtcca acatgtcgaa agcgcagatt atccagcagg ccggtaaactc cgtgctggca 1740
aaagccaacc aggtaccgca gcaggttctg tctctactgc aggtttaa 1788

```

<210> 67
 <211> 1398
 <212> DNA
 <213> Escherichia coli

```

<400> 67
aacaaatctc agtcttctct tagctctgct attgagcgtc tgtcttctgg tctgcgtatt 60
aacagcgcaa aagacgatgc agcaggtcag gcgattgcta accgttttac ggcaaatatt 120
aaagggtotga cccaggtctc ccgtaacgca aatgatggta tttctgttgc gcagaccact 180
gaagggtgcgc tgaatgaaat taacaacaac ctgcagcgta ttcgtgaact ttctgttcag 240
gcaactaacg gtactaactc tgacagtgc ctgacctcca tccagtccga aatccagcag 300
cgtctgagtg aaattgaccg tgtttctggg cagactcagt ttaacggcgt taaagtgtcg 360
gcttctgatac aggatatgac tattcaggtt ggtgcaaacg acggcgaaac aattactatt 420
aaactgcagg aaattaattc cgacacactg ggattatctg gttttggtat taaagatcct 480
actaaattaa aagccgcaac ggctgaaaca acctattttg gatcgacagt taagcttgct 540
gacgctaata cacttgatgc agatattaca gctacagtta aaggcactac gactccgggc 600
caacgtgacg gtaatattat gtctgatgct aacggtaagt tgtacgttaa agttgccggt 660
tcagataaac ccgctgaaaa tgggtattat gaagttactg tggaggatga tccgacatct 720
cctgatgcag gtaagctgaa gctgggggct ctagcgggta cccagcctca agctggtaat 780
ttaaaggaa gtcacaacgg taaagggaa ggggctattg atgttcagtt gggtagctgat 840
accgcaaccg cttctatcac aggtgcacaaa ctctttaagt tagaagacgc caatggcaaa 900
gatactgggt catttgcggt gattggtgat gacggtaaac agtatgcagc gaatgttgat 960
cagaaaacag gagcagtttc cgttaaaaca atgtcttaca ctgatgctga cgggtgcaaa 1020
cacgacaatg ttaaaagtga actgggtgga agcgatggca aaaccgaagt tgtaactgca 1080
accgatggca aaacttacag tgtagtgat ttacaaggta agagcctgaa aactgattct 1140
attgcagcaa tttctacgca gaaaacagaa gatcctttgg ctgctatcga taaagcactg 1200
tctcaggttg actcgttgcg ttctaacctt ggtgcaattc aaaatcggtt cgactctgcc 1260
atcaccaacc ttggcaacac cgtaaacaac ctgtcttctg cccgtagccg tatcgaaagt 1320
gctgactacg cgaccgaagt gtctaacatg tctcgtgcgc agatcctgca acaagcgggt 1380
acctctgttc tggcgacg

```

<210> 68
 <211> 1479
 <212> DNA
 <213> Escherichia coli

```

<400> 68
aacaaatctc agtcttctct gagctccgcc attgaacgtc tctcttctgg cctgcgtatt 60
aacagtgccta aagatgacgc agcaggtcag gcgattgcta accgttttac agcaaatatt 120
aaagggtctga ctcaggcttc ccgtaacgcg aatgatggta tttctgttgc gcagaccact 180
gaagggtgcgc tttctgaaat caacaataac ttacagcgta ttcgtgaatt gtcagtacag 240
gccactaatg gtacaaactc tgactccgac ctgaattcaa ttcaggatga aattacacaa 300
cgcttagtg aaattgatcg tgtttctaac cagacacaa ttaatgggtg aaaagttctg 360
gcttctgatac agactatgaa aattcaagta ggtgcgaacg atggtgaaac cattgagatt 420
gcccttgata aaattgatgc taaaaccttg ggccttgata actttagcgt agcaccagga 480
aaagttccaa tgcctctgct ggttgcaact aagagcgaag ccgctcctga cttaactaag 540
gtaaatgcaa ctgatggtag tgtgggaggt gctaaagcat tcggtagcaa ttataaaaaat 600
gctgatggtt aaacttattt tggtagcggg aatgtacaag atcaaaagga tacaactgat 660
gcgaccggta ctgcaggaaac aaaagtttat caagtacagg tggaggggca gacttatttt 720
ggttggtcaag ataataatac caacacgaac ggttttacat tattgaaaca aaactctaca 780
ggttatgaaa aagttcaggt ggggtggtgaa gatgttcagt tagcaaaact tgggtggtcgt 840

```

[illegible]

WO 99/61458

PCT/AU99/00385

- 1 -

SEQUENCE LISTING PART

<110> THE UNIVERSITY OF SYDNEY

<120> ANTIGENS AND THEIR DETECTION

<130> REEVES

<140>

<141>

<160> 68

<170> PatentIn Ver. 2.0

<210> 1

<211> 1773

<212> DNA

<213> Escherichia coli

<400> 1

atgCGacgta tagaacgaat accgggggta tggcgtaag cggggcaaag tttacgattt 60
atTTTTtggc ttaatgacac gaacagcaac gaggaagggg agtatttcga ccgctagaaa 120
aaaattctaa aggttgtgag tgaccagacg ataacagggt tgacggcgac gaagccgaag 180
ggtggaagcc caatacttaa accgtagact tgaaaacagg aaaatgaatc atggcacaag 240
tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag aaccagtctg 300
cgctgtcgac ttctatcgag cgctctctt ctgggtctgag cattaacagc gctaaagatg 360
acgctgcggg ccaagcgatt gctaaccgct tcaattctaa catcaaagggt ctgactcagg 420
ccgcacgtaa cgccaacgac ggtattttctc tggcgcgac cactgaaggc gcaactgtctg 480
aaatcaacaa caacttgacg cgtgttcgtg aactgaccgt tcaggccact accggtacta 540
actctgatcc tgacctgtct tcaatacagg acgaaatcaa atcccgctctc gatgaaattg 600
accgcgtatc cggtcagact cagttcaacg gcgttaattg tctttccaaa gatggttcaa 660
tgaaaattca ggttggtgag aatgatgggc aaactatctc catcgatctg aagaaaattg 720
attcttcaac tttggggctg aatggcttct cagtttctaa aaactctctt aatgtcagca 780
atgctatcac atctatcccg caagccgcta gcaatgaacc tgttgatggt aacttcgggtg 840
atactgatga gtctgcagca atcgagcca aattgggggt ttccgatacg tcaagcctgt 900
cgctgcacaa catccttgat aaagatggta aggcaacagc tgattatggt gttcagtcag 960
gtaaagactt ctatgctgct tctgttaatt ccgcttcagg taaagtaacc ttaaaccacca 1020
ttgatgttac ttatgatgat tatgcgaacg gtgttgacga tgccaagcaa acaggtcagc 1080
tgatcaaagt ttcagcagat aaagacggcg cagctcaagg ttttgtcaca cttcaaggca 1140
aaaactatct tgctgggtgat gcggcagaca ttcttaagaa tggagcaaca gctcttaagt 1200
taactgatct gaatttaagt gatgttactg atactaatgg taaggtaacc acaactgcga 1260
ctgagcaatt tgaaggtgct tcaactgagg atccgctggc gcttctggat aaagctattg 1320
catcagtcga caaattccgg tcttctctag gtgccgtgca gaaccgtctc gattccgcta 1380
tcaccaacct gaacaacacc accaccaacc tgtctgaagc gcagtcccgat attcaggacg 1440
ccgactatgc gaccgaagtg tccaacatgt cgaaagcgca gatcatccag caggcaggta 1500
actccgtgct gtctaaagcg aaccagggtac cgcagcaagt tctgtcactg ttacaaggct 1560
aatggcctta acctgcctga ccccgccacc ggggggggtt tttctgtccg caatttaccg 1620

ataacccccca aataaccacct catttcacccc actaatcgtc cgattaaaaa ccctgcagaa 1680
acggataatc atgccgataa ctcatataac gcagggctgt ttatcgtgaa ttcactctat 1740
accgctgaag gtgtaatgga taaacactcg ctg 1773

<210> 2

<211> 500

<212> DNA

<213> Escherichia coli

<400> 2

aacagcctct cgctgacac tcagaacaac atcaacaaaa accagtcttc aatgtctact 60
gccattgagc gtctgtcttc cggctctgct atcaacagcg caaaagatga cgctgctggc 120
caggcgattg ccaaccgctt cacctctaac atcaaaggct tgactcaggc agctcgtaac 180
gccaacgacg gtatctccgt tgcacagacc actgaaggcg cactgtctga aatcaacaac 240
aacctgcagc gtatccgtga gctgactggt cagtcttcta cgggtactaa ctctgaatcc 300
gatctgaact caatccagga cgaaattaaa tcccgctctgg acgaaattga ccgcgtatcc 360
ggtcagaccc agttcaacgg cgtgaacgtg ctggcaaaaag acggctccat gaaaattcag 420
gttggcgcgga acgatggtga aaccatcacc atcgacctga aaaaaattga ctcttctact 480
ttaaacctga ctgggtttaa 500

<210> 3

<211> 500

<212> DNA

<213> Escherichia coli

<400> 3

ctcagtatgc tgtcacccgc agtacagggt ccgtaactta cgatccagat acagatcctg 60
ccgcgactgg tgatattggt tctgcttatg ttgatgatgc aggtacattg acaactgatg 120
caaacaaaac tgtaaaatat tatgccaca ctaatggtag cgtcacgaac gacagtgggt 180
cagctattta cgcaactgaa gcgggcaaat tgactactga agcgtctaca gctgctgaaa 240
ctaccgctaa cccactgaaa gccctggacg atgcaatcag ccagatcgac aaattccggt 300
cttctctggg tgctgtacag aaccgtctgg attctgcggt aaccaacctg aacaacacca 360
ccaccaacct gtctgaagcg cagtcccgtg ttcaggacgc cgactatgcg accgaagtgt 420
caaatatgtc taaagcgcag atcatccagc aggcgggtaa ctccgtgttg gctaaagcta 480
accaggttcc tcagcaggtt 500

<210> 4

<211> 399

<212> DNA

<213> Escherichia coli

<400> 4

agcctgtcgc tgttgacca gaataacctg aacaaatctc agtcttctct gagctccgcc 60
attgagcgtc tctcttctgg cctgcgtatt aacagtgcta aagatgacgc agcaggtcag 120
gcgattgcta accgttttac agcaaatatt aaaggctctga ctcaggcttc ccgtaacgcg 180
aatgatggta tttctgttgc gcagaccact gaaggcgcgc tgaatgaaat taacaacaac 240
ctgcagcgtg tacgtgaact gactgttcag gcaactaacg gtactaactc tgacagcgat 300
ctttcttcta tccaggctga aattactcaa cgtctggaag aaattgaccg tgtatctgag 360

caaactcagt ttaacggcgt gaaagtcctt gctgaaaat

399

<210> 5

<211> 417

<212> DNA

<213> Escherichia coli

<400> 5

gcacgttagt tgttaacggg gcaacttacg atgttagtgc agatggtaaa acgataacgg 60
agactgcttc tggtaacaat aaagtcacgt atctgagcaa atcagaaggt ggtagcccca 120
ttctggtaaa cgaagatgca gcaaaatcgt tgcaatctac caccaaccgg ctcgaaacta 180
tcgacaaaagc attgggctaaa gttgacaatc tgcgttctga cctcgggtgca gtacaaaacc 240
gtttcgcactc tgctatcacc aaccttggca acaccgtaaa caacctgtct tctgcccgtta 300
gccgtatcga agatgctgac tacgcgaccg aagtgtctaa catgtctcgt gcgcagatcc 360
tgcaacaagc gggtagctct gttctggcgc aggctaacca gaccacgcag aacgtac 417

<210> 6

<211> 950

<212> DNA

<213> Escherichia coli

<400> 6

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgtatt 60
aacagcgcta aagatgacgc cgcggggccag gcgattgcta accgctttac ttctaacatc 120
aaaggctctga ctacggccgc acgtaacgcc aacgacggta tttctctggc gcagacggct 180
gaaggcgcgc tgcagagat taacaacaac ttgcagcgta ttcgtgaact gaccgttcag 240
gcctctaccg gcacgaactc tgattccgac ctgtcttcta ttcaggacga aatcaaacc 300
cgtcttgatg aaattgaccg tgtatctggt cagacccagt tcaacgggtg gaacgtgctg 360
tcgaaaaacg attcgatgaa gattcagatt ggtgccaatg ataaccagac gatcagcatt 420
ggcttgcaac aaatcgacag taccactttg aatctgaaag gatttacgtg gtccggcatg 480
gcggatttca gcgcggcgaa actgacggct gctgatggta cagcaattgc tgctgcggat 540
gtcaaggatg ctgggggtaa acaagtcaat ttactgtctt acactgacac cgcgtctaac 600
agtactaaat atgcggctcg tgattctgca accggtaaat acatggaagc cactgtagtc 660
attaccggta cggcggcgcc ggtaactgtt ggtgcagcgg aagtggcggg agccgctaca 720
gccgatccgt taaaagcact ggatgccgca atcgctaaag tcgacaaatt ccgctctccc 780
ctcgggtgccg ttcaaaaccg tctggattct gcggtcacca acctgaacaa caccaccacc 840
aacctgtctg aagcgcagtc ccgtattcag gacgcgact atgcgaccga agtgtccaac 900
atgtcgaaag cgcagattat ccagcaggcg ggcaactccg tgctgtctaa 950

<210> 7

<211> 1212

<212> DNA

<213> Escherichia coli

<400> 7

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgtatt 60
aacagcgcta aagatgacgc cgcggggccag gcgattgcta accgcttcac ttctaacatc 120
aaaggctctga ctacggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180

gaaggcgcgc tgtctgaaat caacaacaac ttgcagcgtg tgcgtgagtt gaccgttcag 240
gcgacgaccg ggactaactc tgattctgac ctgtcttcta ttcaggacga aatcaaatcc 300
cgtctggatg aaattgatcg cgtttccggt cagaccaggt tcaacggcgt gaatgtgctg 360
gcgaaagatg gttogatgaa gattcaggtt ggcgcgaatg atgggcagac tattagcatt 420
gatttgacga agattgactc ttctacatta ggactgaacg gtttctccgt ttcgggtcag 480
tacttaacg ttagtgattc cactactcaa attaccggtg ccgccgggac aaaacctgtt 540
ggtgttgatt tcaactgctg tgcgaaagat ctgactactg cgacaggtaa aacagtcgat 600
gtttctagcc tgacgttaca caacactctg gatgcgaaag gggctgctac atcacagttc 660
gtcgttcaat ccggcaatga tttctactcc gcgtcgatta atcatacaga cggcaaagtc 720
acgttgaata aagccgatgt cgaatacaca gacaccgata atggactaac gactgcggct 780
actcagaaag atcaactgat taaagttgcc gctgactctg acggctcggc tgcgggatat 840
gtaacattcc aaggtaaaaa ctacgctaca acggtttcaa cggcacttga tgataatact 900
gcggcaaaag caacagataa taaagttggt gttgaattat caacagcaaa accgactgca 960
cagttctcag gggcttcttc tgctgatcca ctggcacttt tagacaaagc tattgcacag 1020
gttgatactt tccgctcttc cctcgggtgcg gtgcaaaacc gtctggattc cgcagtaacc 1080
aacctgaaca acaccaccac caacctgtct gaagcgcagt cccgtattca ggacgccgac 1140
tatgctacag aagtgtccaa catgtcgaaa gcgcagatca tccagcaggc aggttaactcg 1200
gtgctgtcca aa 1212

<210> 8

<211> 1647

<212> DNA

<213> Escherichia coli

<400> 8

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcccgagg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtccg aaattaacaa caacttacag cgtattcgtg aactgacggt tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaaaattca ggttgggtgcg aatgacggcc agactatcac tattgatctg 480
aagaaaattg actctgatac gctggggctg aatgggttta atgtgaacgg caaaggggaa 540
acggctaata cggcagcaac cctgaaagat atgtctggat tcacagctgc ggcggcacca 600
gggggaactg ttggtgtaac tcaatatact gacaaatcgg ctgtagcaag tagcgtagat 660
attctaaatg ctggtgctgg cgcatatgga aataaagtta caactagcgc cgatgttggt 720
tttggtacac cagccgctgc tgtaacctat acctacaata aagacactaa ttcatttcc 780
gccgcttctg atgatatttc cagcgctaac ctggctgctt tctcaatcc tcaggccgga 840
gatacgacta aagctacagt tacaattggt ggcaaagatc aagatgtaaa catcgataaa 900
tccggttaatt taactgctgc tgatgatggc gcagtacttt atatggatgc taccggtaac 960
ttaactaaaa ataattgctg tggtgataca caagctactt tggctaaact tgctactgct 1020
actggtgcta aagccgcgac catccaaact gataaaggaa cattcaccag tgacggtaca 1080
gcgtttgatg gtgcatcaat gtccattgat accaatacat ttgcaaatgc agtaaaaaat 1140
gacacttata ctgccactgt aggtgctaag acttatagcg taacaacagg ttctgctgct 1200
gcagacaccg cttatatgag caatgggggt ctcaagtata ctccgccaac ttactatgca 1260
caagctgatg gaagtatcac aactactgag gatgcggctg ccggtaaact ggtctacaaa 1320
ggttccgatg gtaagttaac aacggatagc actagcaaa gagaatcaac atcagatccg 1380

ctggcagctc ttgacgacgc tatcagccag atcgacaaat tccgctcctc cctgggtgcg 1440
gtgcaaaacc gtctggattc cgcagtgacc aacctgaaca acaccactac caacctgtct 1500
gaagcgcagt cccgtattca ggacgcgcag tatgcgaccg aagtgtccaa catgtcgaaa 1560
gcgcagatta tccagcaggc cggtaactcc gtgctggcaa aagctaacca gggtccgcag 1620
cagggtctgt ctctgctgca gggtttaa 1647

<210> 9

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 9

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac caccgaaggc 240
gcgctgtctg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggcttct 300
accggaacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgctctt 360
gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaaaattca ggttgggtgcg aatgacggtg aaactatcac tatcgacctg 480
aagaaaatcg attctgatac tctgggtctg aatggtttta acgtaaattg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaaccacc 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcacc gttggcggcg tagattatac ttacaacgct 720
aaatctgggt attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaactctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct taaagccgcg 1020
agcgaaggta gtgacggtgc ttctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctgggtgggag tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtggtatta ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtactg ctgtaaattg gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
atccagcagg ccggtaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa 1758

<210> 10

<211> 1383

<212> DNA

<213> Escherichia coli

<400> 10

aacaaatctc agtcttctct tagctctgct attgagcgtc tgtcttctgg tctgcgtatt 60
aacagcgcaa aagacgatgc agcaggtcag gcgattgcta accgttttac ggcaaatatt 120
aaaggtctga cccaggtctc ccgtaacgca aatgatggta tttctgttgc gcagaccact 180
gaaggtgccc tgaatgaaat taacaacaac ctgcagcgta ttcgtgaact ttctgttcag 240
gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggctga aattactcaa 300
cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360
gctgaaaata atgaaatgaa aattcaggtt ggtgctaatt atggtgaaac catcactatc 420
aatctggcaa aaattgatgc gaaaactctc ggccctggacg gttttaatat cgatggcgcg 480
cagaaagcaa caggcagtga cctgatttct aaatttaaag cgacaggtag tgataattat 540
gatgttggcg gtaaaactta taccgtgaat gtggagagcg gcgcgggttaa gaatgatgct 600
aataaagatg tttttgtaag cgcagctgat ggatcgctga cgaccagtag tgatactaaa 660
gtatccgggtg aaagtattga tgcaacagaa ctagcgaaac ttgcaataaa attagctgac 720
aaaggctcca ttgaatacaa gggcattaca tttactaaca aactggcgcg agagcttgat 780
gctaattggtg aaggtgtttt gaccgcaaat attgatggtc aagatgttca atttactatt 840
gacagtaatg caccacggg tgccggcgca acaataacta cagacacagc tgtttacaaa 900
aacagtggcg gccagttcac cactacaaaa gtggaaaata aagccgcaac actctctgat 960
ctggatctta atgcagccaa gaaaacagggt agcacttttag ttgtaaatgg cgccacctac 1020
aatgtcagcg cagatggtaa aacggtaact gatactactc ctggtgcccc taaagtgatg 1080
tatctgagca aatcagaagg tggtagcccc attctggtaa acgaagatgc agcaaaatcg 1140
ttgcaatcta ccaccaaccc gtcgaaaact atcgacaagg cattggctaa agttgacaat 1200
ctgcgttctg acctcggtgc agtacaaaac cgtttcgact ctgccatcac caaccttggc 1260
aacaccgtaa acaacctgtc ttctgcccgt agccgtatcg aagatgctga ctacgcgacc 1320
gaagtgtcta acatgtctcg tgcgcagatc ctgcaacaag cgggtacctc tgttctggcg 1380
cag 1383

<210> 11

<211> 2013

<212> DNA

<213> Escherichia coli

<400> 11

atggcacaaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgtctgcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttccg ttgcacagac cactgaaggc 240
gcgctgtccg aaattaacaa caacttacag cgtattcgtg aactgacggg tcaggcttct 300
accgggacta actccgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc cgccagacc cagttcaacg gcgtgaacgt gctgtccaaa 420
gatggctcga tgaaaattca ggtcggcgcg aacgatggcg aaacgattac tattgatctg 480
aagaaaattg actctgatac gctgaatctg gctgggttta acgttaacgg taaaggttct 540
gtagcgaata cagctgcgac aagcgacgat ttaaaactgg ctggtttcac taagggcacc 600
acagatacca atggcggtgac cgcgtatata aacacaatta gtaatgacaa agccaaagct 660
tccgatctgt tagctaatat caccgatgga tcagtgatca ctgggggagg ggcaaacgct 720
tttggcggtg ctgcaaagaa tggttacacc tatgatgcag caagtaaatc ttatagtttt 780
gctgcagatg gtgccgattc agcgaagacg ttaagcatca ttaatccaaa caccgggtgat 840
tcgtcgcagg cgacagtgc tattggtggt aaagagcaga aagttaatat ttcccaggat 900
ggaaaaatta ctgcggcaga tgataatgcg acgctgtatt tagataaaca gggaaacttg 960

acaaaaacga atgcaggtaa cgataccgca ggcacttggg atggtttaaat ttccaacagc 1020
gattctaccg gtgcgggtcc agttgggggtt gcaactacaa ttacaattac ttctgggtaca 1080
gcttcgggaa tgtctgttca gtccgcagga gcaggaattc agacctcaac aaattctcag 1140
attcttgcag gtgggtgcatt tgccggctaag gtaagtattg agggaggcgc tgctacagac 1200
atcttggtag caagtaatgg aacataaca gcggctgatg gtagtgcact ttatcttgat 1260
gcgactactg gtggattcac tacaacggct ggaggaaata cagctgcttc gttagataat 1320
ttaattgcta acagtaagga tgctacctta accgtaactt caggtaccgg ccagaacact 1380
gtttatagca caacaggaag tggcgctcag ttcaccagtt tagcaaaagt agacacagtc 1440
aatgtcacca acgcacatgt cagtgcgaa ggtatggcaa atctgacaaa aagcaatttt 1500
accattgata tggcggtac aggtacagta acttacacag tttccaatgg ggatgtgaaa 1560
gctgctgcaa atgctgatgt ttatgtcgaa gatgggtgcac tttcagccaa tgctacaaaa 1620
gatgtaacct actttgaaca aaaaaatggg gctattacca acagcaccgg tgggtaccatc 1680
tatgaaacag ctgatggtaa gttaacaaca gaagctacta ctgcatccag ttccaccgcc 1740
gatccccctga aagctctgga cgaagccatc agctccatcg acaaattccg ctctccctc 1800
gggtgcggtgc aaaaccgtct ggattccgcg gtcaccaacc tgaacaacac cactaccaac 1860
ctgtccgaag cgcagtcccg tattcaggac gccgactatg cgaccgaagt gtccaacatg 1920
tcgaaagcgc agatcatcca gcaggccggt aactccgtgc tggcaaaagc taaccaggta 1980
ccgcagcagg ttctgtctct gctgcagggt taa 2013

<210> 12

<211> 1263

<212> DNA

<213> Escherichia coli

<400> 12

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcggtc tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtccg aaattaacaa caacttacag cgtgtgcgtg agctgactgt tcaggcgacc 300
accggtacta actctgagtc tgacctgtct tctatccagg acgaaatcaa atctcgctg 360
gaagagattg atcgtgtttc aagtcagact caatttaacg gcgtgaatgt tttggctaaa 420
gatgggaaaa tgaacattca gggtggggca aatgatggac agactatcac tattgatctg 480
aaaaagatcg attcatctac actaaacctc tccagttttg atgctacaaa cttgggcacc 540
agtgttaaag atggggccac catcaataag caagtggcag taggtgctgg cgactttaaa 600
gataaagctt caggatcggt aggtacccta aaattagttg agaaagacgg taagtactat 660
gtaaagtaca ctaaaagtag taagtactac gatgccgaag tagatactag taagggtaaa 720
attaacttca actctacaaa tgaagtggga actactccta ctgcagcgac ggaagtaact 780
actgttgccc gcgatgtaaa attggatgct tctgcactta aagccaacca atcgcttgct 840
gtgtataaag ataaaagcgg caatgatgct tatatcattc agaccaaaga tgtaacaact 900
aatcaatcaa ctttcaatgc cgctaatac agtgatgctg gtgttttatc tattgggtga 960
tctacaaccg cgccaagcaa tttaacagct aaccgcgtta aggtcttga tgatgcaatt 1020
gcatctgttg ataaattccg ctcttctctc ggtgccgttc agaaccgtct ggattctgcc 1080
attgccaacc tgaacaacac cactaccaac ctgtctgaag cgcagtcccg tattcaggac 1140
gctgactatg cgaccgaagt gtccaacatg tcgaaagcgc agattatcca gcaggccggt 1200
aactccgtgc tggcaaaagc caaccaggta ccgcagcagg ttctgtctct gctgcagggt 1260
taa 1263

<210> 13

<211> 1368

<212> DNA

<213> Escherichia coli

<400> 13

aacaaatctc agtcttctct gagctccgcc attgaacgtc tctcttctgg cctgcgtatt 60
aacagtgcta aagatgacgc agcaggtcag gcgattgcta accgttttac agcaaattatt 120
aaaggtctga ctcaggcttc ccgtaacgcg aatgatggta tttctgttgc gcagaccact 180
gaaggtgctc tgaatgaaat taacaacaac ctgcagcgtg tacgtgaact gactgttcag 240
gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggctga aattactcaa 300
cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360
gctgaaaata atgaaatgaa aattcagggtt ggtgctaata atggtgaaac catcactatc 420
aatctggcaa aaattgatgc gaaaactctc ggcctggacg gttttaatat cgatggcgcg 480
cagaaagcaa ctggcagtgta cctgatttct aaatttaaag cgacaggtag tgataactat 540
gatgttggcg gtgatgctta tactgttaac gtagatagcg gagctgttaa agatactaca 600
gggaatgata tttttgttag tgcagcagat gggtcactga caactaaatc tgacacaaac 660
atagctggta cagggattga tgctacagca ctgcagcag cggttaagaa taaagcacag 720
aatgataaat tcacgtttta tggagttgaa ttcacaacaa caactgcagc ggatggcaat 780
gggaatggtg tatattctgc agaaattgat ggtaagtcag tgacatttac tgtgacagat 840
gctgacaaaa aagcttcttt gattacgagt gagacagttt acaaaaatag cgctggcctt 900
tatacgacaa ccaaagttga taacaaggct gccacacttt ccgatcttga tctcaatgca 960
gctaagaaaa caggaagcac gttagttggt aacggtgcaa cttacgatgt tagtgcagat 1020
ggtaaaacga taacggagac tgcttctggt aacaataaag tcatgtatct gagcaaataca 1080
gaaggtggta gcccgattct ggtaaacgaa gatgcagcaa aatcggttga atctaccacc 1140
aaccgcctcg aaactatcga caaagcattg gctaaagttg acaatctgcy ttctgacctc 1200
ggtgcagtag aaaaccgttt cgactctgct atcaccaacc ttggcaacac cgtaaacaac 1260
ctgtcttctg cccgtagccg tatcgaagat gctgactacg cgaccgaagt gtctaactatg 1320
tctcgtgcgc agatcctgca acaagcgggt acctctgttc tggcgag 1368

<210> 14

<211> 1788

<212> DNA

<213> Escherichia coli

<400> 14

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcy tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt tcacctctaa cattaaggc 180
ctgactcagg cggcccgtaa cgccaacgac ggtatctccg ttgcgcagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggcttct 300
accgggacta actccgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc tggccagacc cagttcaacg gcgtgaacgt actggcgaaa 420
gacggttcaa tgaaaattca ggttggtgcy aatgacggcc agactatcac gattgatctg 480
aagaaaattg actcagatac gctggggctg aatggtttta acgtgaatgg ttccggtacg 540
atagccaata aagcggcgac cattagcgac ctgacagcag cgaaaatgga tgctgcaact 600
aatactataa ctacaacaaa taatgcgctg actgcatcaa aggcgcttga tcaactgaaa 660
gatggtgaca ctgttactat caaagcagat gctgctcaaa ctgccacggt ttatacatatc 720

aatgcatcag ctggttaactt ctcattcagt aatgtatcga ataatacttc agcaaaagca 780
ggtgatgtag cagctagcct tctcccgccg gctgggcaaa ctgctagtgg tgtttataaa 840
gcagcaagcg gtgaagtga ctttgatgtt gatgcgaatg gtaaaatcac aatcggagga 900
cagaaagcat atttaactag tgatggtaac ttaactacaa acgatgctgg tggcgcgact 960
gcggtacgc ttgatggttt attcaagaaa gctgggtgatg gtcaatcaat cgggtttaag 1020
aagactgcat cagtcacgat ggggggaaca acttataact ttaaaacggg tgctgatgct 1080
gatgctgcaa ctgctaacgc aggggtatcg ttcactgata cagctagcaa agaaaccgtt 1140
ttaaataaag tggctacagc taaacaaggc aaagcagttg cagctgacgg tgatacatcc 1200
gcaacaatta cctataaatc tggcggttcag acgtatcagg ctgtatttgc cgcaggtgac 1260
ggtactgcta gcgcaaaata tgccgataaa gctgacgttt ctaatgcaac agcaacatac 1320
actgatgctg atgggtgaaat gactacaatt gggttcataca ccacgaagta ttcaatcgat 1380
gctaacaacg gcaaggtaac tgttgattct ggaactggta cgggtaaata tgcgccgaaa 1440
gtaggggctg aagtatatgt tagtgctaatt ggtactttta caacagatgc aactagcgaa 1500
ggcacagtaa caaaagatcc actgaaagct ctggatgaag ctatcagctc catcgacaaa 1560
ttccgttctt ccttgggtgc tatccagaac cgtctggatt ccgcagtcac caacctgaac 1620
aacaccacta ccaacctgtc cgaagcgcag tcccgtatc aggacgccga ctatgcgacc 1680
gaagtgtcca acatgtcgaa agcgcagatc attcagcagg ccggtaaact cgtgctggca 1740
aaagccaacc aggtaccgca gcaggttctg tctctgctgc agggttaa 1788

<210> 15

<211> 1653

<212> DNA

<213> Escherichia coli

<400> 15

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gogaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttccg ttgcgcagac cactgaagggt 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg agctgacggg tcaggcttct 300
accgggacta actccgattc tgacctggac tccatccagg acgaaatcaa gtctcgtctg 360
gacgaaattg accgcgtatc oggtcagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacgggtcga tgaaaattca ggttgggtgcg aatgacggcc agactatcac gattgatctg 480
aagaaaattg actcagatac gctggggctg agtgggttta atgtgaatgg tggcggggct 540
gttgctaaca ctgctgcata taaagctgac ttggtagctg ctaatgcaac tgtggtaggc 600
aacaaatata ctgtgagtgc gggttacgat gctgctaaag cgtctgattt gctggctgga 660
gttagtgatg gtgatactgt tcaggcaacc attaataacg gcttcggaac ggccgctagt 720
gcaacgaatt acaagtatga cagtgaagt aagtcttact cttttgatac cacaacggct 780
tcagctgccg atgttcagaa atatttgacc ccgggcgttg gtgataccgc taagggcact 840
attactatcg atggttctgc acaggatgtt cagatcagca gtgatggtaa aattacgtca 900
agcaatggag ataaacttta cattgataca actgggcgct taacgaaaaa cggcttttagt 960
gcttctttga ctgaggctag tctgtccaca cttgcagcca ataataccaa agcgacaacc 1020
attgacattg gcggtacctc tatctctttt accggttaata gtactacgcc gaacactatt 1080
acttattcag taacaggtgc aaaagttgat caggcagctt tcgataaagc tgtatcaacc 1140
tctggaaacg atgttgattt cactaccgca ggttatagcg tcgacggcgc aactggcgct 1200
gtaacaaaag gtgttgctcc ggtttatatt gataacaacg gggcggtgac cacatctgat 1260
actgtagatt tttatctaca ggatgatggg tcagtgacta acggcagcgg taaggcagtt 1320
tataaagatg ctgacggtaa attgacgaca gatgctgaaa ctaaagctgc aaccaccgcc 1380

gatccctga aagctctgga cgaagccatc agtccatcg acaaattccg ctctccctc 1440
ggtgcggtgc agaaccgtct ggattccgcy gtcaccaacc tgaacaacac cactaccaac 1500
ctgtctgaag cgcagtcctc tattcaggac gctgactatg cgaccgaagt atccaacatg 1560
tcgaaagcgc agatcatcca gcaggccggt aactccgtgc tggcaaaagc taaccaggta 1620
ccacagcagg ttctgtctct gctgcagggt taa 1653

<210> 16

<211> 1689

<212> DNA

<213> Escherichia coli

<400> 16

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgtctgtagg ttctatcgag cgtctgtctt ctggcttgcy tattaacagc 120
gcgaaggatg acgcgcgagg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtgtgcgtg aactgaccgt tcaggcaacc 300
accggtacca actcccagtc tgacctggac tctatccagg acgaaattaa atcccgtctg 360
gacgaaattg atcgcgtatc cggtcagacc cagttcaacg gcgtgaacgt gctggcaaaa 420
gacggttcca tgaaaattca ggttggcgcg aacgatggcc agaccatcac tatcgacctg 480
aagaagattg actcttctac cttgaacctg acaggtttta acgttaacgg ttctggttct 540
gtggcgaata ctgcagcaac taaagctgat ttaaccgtg ctcaactctc tgcaccgggt 600
gcagcagacg caaatggtac agttacttat actgtcagtg ctggttataa agaatccact 660
gctgcagatg ttattgctag catcaaagac ggagtgctc cgacttctgc aattactgca 720
accattaata atggcttcgg tgattccagt gcgtgactt ccaatgacta tacttatgac 780
ccagcaaaag gcgacttcac ttacgacgta gcttcaagcg ccaataatac tgctgccag 840
gttcagtcct tctgacgcc gaaagcagg gataccgcaa atctgaaagt aaccgttggt 900
acgacatcgg ttgatgtcgt tctggccagt gatggtaaga ttacagcaaa agatggttct 960
gcattatata tcgacagtac aggtaacctg actcagaaca gtgctggctt gacctctgct 1020
aaactggcta ctctgactgg ccttcagggc tctggtgttg ctccaacat cactactgaa 1080
gatggcacta atattgatat tgctgctaac ggtaatatg gtctgaccgg tgttcgtatc 1140
agtgtgatt ctctgcagtc agcgactaaa tctacgggct ttactgttgg tactggcgct 1200
acaggtctga ccgtaggtac tgatggtaaa gtgactatcg gcgggactac tgctcagtc 1260
tacaccagca aagatggttc cctgactact gataacacca ctaaactgta tctgcagaaa 1320
gatggctctg taaccaacgg ttcaggtaaa gcggtctatg tagaagcgga tgggtatttc 1380
actaccgacg ctgcaaccaa agccgcaacc accaccgatc cgctgaaagc cctggatgag 1440
gcaatcagcc agatcgataa gtcccggtca tccctgggtg ctatccagaa ccgtctggat 1500
tccgcgggtca ccaacctgaa caacaccact accaactgt ctgaagcgca gtcccgtatt 1560
caggacgccg actatgcgac cgaagtgtcc aacatgtcga aagcgcagat cattcagcag 1620
gccggttaact ccgtgctggc aaaagccaac cagggtaccg aacaggttct gtctctgctg 1680
cagggtctaa 1689

<210> 17

<211> 915

<212> DNA

<213> Escherichia coli

<400> 17


```

gcgctgtcga cttctatcga ggcctctctt tctggtctgc gtattaacag cgctaaagat 60
gacgctgcgg gccaggcgat tgctaaccgc ttcacttcta acatcaaagg tctgactcag 120
gccgcacgta acgccaacga cggattttct ctggcgcgaga cggctgaagg cgcgctgtca 180
gagattaaca acaacttgca gcgtattcgt gaactgaccg ttcaggcctc taccggcacg 240
aactctgatt ccgacctgtc ttctattcag gacgaaatca aatcccgctc tgatgaaatt 300
gaccgtgtat ctggtcagac ccagttcaac ggtgtgaacg tgctgtcgaa aaacgattcg 360
atgaagattc agattggtgc caatgataac cagacgatca gcattggctt gcaacaaatc 420
gacagtacca ctttgaatct gaaaggattt accgtgtccg gcatggcgga tttcagcgcg 480
gcgaaactga cggctgctga tggtagacga attgctgctg cggatgtcaa ggatgctggg 540
ggtaaacaag tcaatttact gtcttacact gacaccgct ctaacagtac taaatatgcg 600
gtcgttgatt ctgcaaccgg taaatacatg gcagccactg tagtcattac cagtacggcg 660
gcggcggtaa ctgttggtgc aacggaagtg gcgggagccg ctacagccga accgttaaaa 720
gcaactggat ccgcaatcgc taaagtcgac aaattccgct cctccctcgg tgccgttcaa 780
aaccgtctgg attctgcggt caccaacctg aacaacacca ccaccaacct gtctgaagcg 840
cagtcctcgt ttcaggacgc cgactatgcg accgaagtgt ccaacatgct gaaagcgag 900
attatccagc aggcg

```

915

<210> 18

<211> 1665

<212> DNA

<213> Escherichia coli

<400> 18

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa tattaaggc 180
ctgactcagg ctgcacgtaa cgccaatgac ggtattttctg ttgcacagac cactgaaggc 240
gcgctgtcgg aaatcaacaa caacttacag cgtattcgtg aactgacggt tcaggccact 300
acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctg 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt gctgtccaaa 420
gatgggttcaa tgaaaattca ggtcggcgca aatgatggtg aaaccatcac gattgatctg 480
aagaaaattg actctgatac gctgaatctg gctggtttta acgtgaatgg cgaagggtgaa 540
acagccaata ctgctgcaac acttaaagat atgggttggt taaaactcga taatacggg 600
gtcactacag ctggagttaa tagatatatt gctgacaaag ccgtcgcaag tagcacggat 660
attttgaatg cggtagctgg tgttgatggc agtaaaagtt ccacggaggc agatgttggt 720
tttgggtcag ctgccctgg tacgccagtg gaataactt atcataaaga tactaacaca 780
tatacggctt ctgcttcagt tgatgcgact caactggcgg cattcctgaa tcctgaagcg 840
ggtggtacca ctgctgcaac agtaagtatt ggcaacggtg caacagctca agagcaaaaa 900
gtcattattg ctaaagatgg ttctttaact gctgctgatg acggtgccgc tctctatctt 960
gatgatactg gtaacttaag taaaactaac gcaggcactg atactcaagc taaactgtct 1020
gacttaattg caaacaatgc taatgcaaaa acagtcatta caacagataa aggtacattt 1080
actgtaata cgacaaagtt tgatggggta gatatttctg ttgatgcttc aacgtttgct 1140
aacgcggtta aaaatgagac ttacactgca actggttggtg taactttacc tgcgacatat 1200
acagtcaata atggcactgc tgcacagcg tatttagtcg atggaaaagt gagcaaaact 1260
cctgccgagt attttgctca agctgatggc actattacta gtggtgaaaa tgcggctacc 1320
agtaaagcta tctatgtaag tgccaatggt aacttaacga ctaatacaac tagtgaatct 1380
gaagctacta ccaaccgct ggcagcattg gatgacgcta tcgcgtctat cgacaaattc 1440
cgttcttccc tgggtgctat ccagaaccgt ctggattccg cagtcaccaa cctgaacaac 1500

```

accactacca acctgtctga agcgagtcct cgtatttcagg acgcccacta tgcgaccgaa 1560
gtgtccaaca tgtcgaaagc gcagatcatt cagcaggccg gtaactccgt gctggcaaaa 1620
gccaaaccagg taccgcagca ggttctgtct ctgctgcagg gttaa 1665

<210> 19

<211> 1842

<212> DNA

<213> Escherichia coli

<400> 19

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcccagcgc tcaggcgatt gctaaccggt ttacttctaa cattaagggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgctgcagac cactgaaggc 240
gcgctgtccg aaattaacaa caacttacag cgtattcgtg aactgacggt tcaggcgacg 300
accggaacta actccacctc tgacctggac tccatccagg acgaaatcaa atcccgtctt 360
gacgaaattg accgcgtatc tggctcagacc cagtccaacg gcgtgaacgt gctgtctaaa 420
gatggctcga tgaaaattca ggtcggcgcg aacgatggcg aaacgattac tattgatctg 480
aagaaaattg actctgatac gctgaatctg gctggtttta acgttaacgg taaaggttct 540
gtagcgaata ccgctgcgac tacagataat ctgacattgg ctgggtttac agcgggtact 600
aaagctgctg atggcaccgt aacttatagc aaaaatgtcc agtttgccgc cgcgactgca 660
agcaatgtac tggctgctgc taaagatggc gacgaaatta cgttcgtggt taataacggc 720
acaggtatag ctgcaactgg ggggacttat acttatcata aggactctaa ctcatcacgc 780
tttagcgcga cggctgcac taaagattct ctgttgagca cactggcacc aaacgctggc 840
gatacattta ccgctaaagt gactattggg tctaaatcgc aagaagttaa cgttagcaaa 900
gatggtacga ttacatccag cgatggtaag gcgctgtatt tagatgagaa gggcaacctg 960
acccaaacag gtagtggcac aaccaaagct gcaacctggg ataacctgat ggccaatata 1020
gatactacag gcaaagatgc ctatggtaac tctgcccagc cagctgttgg gacagtaatc 1080
gaagcaaaag gaatgaccat cacttctgct ggtggtaatg ctcagggtgtt aaaagacgcg 1140
gcttataatg ccgcatatgc gacctcaatt actactggta ctccgggtga tgcgggagcc 1200
gcgggagccg ctgcaactgc gggtaatgcc gcggtgggag cgctgggccc aacggcagtt 1260
gataatacca cggcagatgt tgccgatatc tctatctcag cttcgcaaat ggcgagcatc 1320
cttcaggata aagatttcac cttaagtgat ggtagtata cttacaacgt gaccagcaat 1380
gctgtcacta tcaatggcaa agcagcaaac attgatgaca gcggcgcaat cacagaccaa 1440
accagtaaag ttgtcaatta ttctgctcat actaacggta gcgtgactaa cgatacaggc 1500
tccactatct atgcgacaga agatggtagc ctgaccaccg atgcagcaac caaagccgaa 1560
accaccgccc atccccctgaa agctctggac gaagccatca gctccatcga caaattccgc 1620
tctccctcgt gtgcggtgca aaaccgtctg gattccgcgg tcaccaacct gaacaacacc 1680
accaccaacc tgtctgaagc gcagtcccgt attcaggacg ccgactatgc gaccgaagtg 1740
tccaacatgt cgaaagcgca gattatccag caggccggta actccgtgct ggcaaaagct 1800
aaccaggtag cacagcaggt tctgtctctg ctgcagggtt aa 1842

<210> 20

<211> 1731

<212> DNA

<213> Escherichia coli

<400> 20

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttactttctaa cattaaaggc 180
ctgactcagg cggcccgtaa cgccaacgac ggtatttctg ttgctgcagac caccgaaggc 240
gcgctgtccg aaattaacaa caacttacag cgtgtgctg agctgactgt tcaggcgacc 300
accggtacca actcccagtc tgatctggac tctatccagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt gctggcaaaa 420
gacggttcca tgaaaattca ggttgccgag aatgatggcc agaccatcac tatcgacctg 480
aagaagattg actcttctac gttgaaactg actggtttta acgtgaatgg ttctggttct 540
gtggcgaata ctgcggcgac taaagcggat ttggctgctg ctgcaattgg taccctgagg 600
gcagcagatt ctacaggtgc cattgcttac acagtaagtg ctgggctgac taaaactaca 660
gccgcagatg tactgtctag cctcgtgat ggtacgacta ttacagccac aggcgtgaaa 720
aatggctttg ctgcaggagc cacttccaat gcctataaac ttaacaaaga taataatata 780
tttacttatg acacgactgc tacgacagct gagctgcagt cttacctgac tccgaaagcg 840
ggcgacactg caacattcag tgttgaaatt ggtggacta cacaagacgt cgtgctgtcc 900
agtgatggca aactcactgc taaggatggc tctaagcttt acattgatac aactggtaat 960
ttaactcaga atggtggtta taacggtgtt ggaacactcg cggaagcgac tctgagtggg 1020
ttagctctga acaaaaatgg tttaacggct gttaaatcca caattactac agctgataac 1080
acttcgattg tactgaatgg ttcaagcgat ggtactggta atgctggtac tgaaggtagc 1140
attgctgtta caggcgtgt aattagttca gctgctctgc aatctgcaag caaacgact 1200
ggtttcactg ttggtacagt agacacagct ggttatatct ctgtaggtac tgatgggagt 1260
gttcaggcat atgatctgc gacttctggc aacaaagctt cttacacca cactgacggg 1320
acactgacta ctgataaacac cactaaactg tatctgcaga aagatggctc tgtaaccaac 1380
ggttcaggta aagcggctta tgtagaagcg gatggtgatt tcaactaccg cgtgcaacc 1440
aaagccgcaa ccaccaccga tccgctggcc gctctggatg acgcaatcag ccagatcgac 1500
aagttccgtt catccttggg tgctatccag aaccgtctgg attctgcagt caccaacctg 1560
aacaacacca ccaccaacct gtctgaagcg cagtcgccgta ttcaggacgc cgactatgcg 1620
accgaagtgt ccaatatgtc gaaagcgag atcatccagc aggcgggtaa ctccgtgctg 1680
gcaaaagcca accaggtacc gcagcaggtt ctgtctctgc tgcagggtta a 1731

<210> 21

<211> 1380

<212> DNA

<213> Escherichia coli

<400> 21

aacaaatctc agtcttctct gagctccgcc attgaacgtc tctcttctgg cctgcgtatt 60
aacagtgcta aagatgacgc agcaggtcag gcgattgcta accgttttac agcaaatatt 120
aaaggtctga ctcaggcttc ccgtaacgcg aatgatggta tttctgttgc gcagaccact 180
gaaggtgcgc tgaatgaaat taacaacaac ctgcagcgta ttcgtgaact ttctgttcag 240
gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggctga aattactcaa 300
cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360
gctgaaaata atgaaatgaa aattcaggtt ggtgctaatt atggtgaaac catcactatc 420
aatctggcaa aaattgatgc gaaaactctc ggcctggacg gttttaatat cgatggcgcg 480
cagaaagcaa ccggcagtg cctgatttct aaatttaaag cgacaggtac tgataattat 540
caaattaacg gtactgataa ctatactgtt aatgtagata gtggcgtagt acaggataaa 600
gatggcaaac aagtttatgt gagtactgcg gatggttcac ttacgaccag cagtgatact 660
caattcaaga ttgatgcaac taagcttgca gtggctgcta aagatttagc tcaagggaat 720

aagattgtct acgaaggtat cgaatttaca aataccggca ctgtcgctat agatgccaaa 780
ggtaatggta aattaaccgc caatgttgat ggtaaggctg ttgaattcac tatttcgggg 840
agtactgata catcaggtac tagtgcaacc gttgcccta cgacagccct atacaaaaat 900
agtgcagggc aattgactgc aacaaaagtt gaaaataaag cagcgacact atctgatctt 960
gatctgaacg ctgccaaagaa aacaggaagc acggttagttg ttaacggtgc aacttacgat 1020
gttagtgcag atggtaaaac gataacggag actgcttctg gtaacaataa agtcatgtat 1080
ctgagcaaat cagaaggtgg tagcccgatt ctggtaaagc aagatgcagc aaaatcgttg 1140
caatctacca ccaaccgct cgaaactatc gacaaagcat tggctaaagt tgacaatctg 1200
cgttctgacc tcgggtgcagt acaaaaccgt ttcgactctg ccatcaccaa ccttggaac 1260
accgtaaca acctgtcttc tgcccgtagc cgtatcgaag atgctgacta cgcgaccgaa 1320
gtgtctaaca tgtctcgtgc gcagatcctg caacaagcgg gtacctctgt tctggcacag 1380

<210> 22

<211> 1767

<212> DNA

<213> Escherichia coli

<400> 22

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccggt ttacttctaa cattaaaggc 180
ctgactcagg cggcacgtaa cgccaacgac ggtatctctc tggcgagac caccgaaggc 240
gcgctgtctg aaatcaacaa caacttacag cgtgtacgtg aactgaccgt tcaggcaacc 300
accggtacta actccgactc cgacctggct tctattcagg acgaaatcaa atcccgctctg 360
gatgaaattg accgcgtatc tggtcagact cagttcaacg gcgtgaacgt gctggcaaaa 420
gacggttcca tgaaaattca ggtaggtgct aacgacggcc agactatcac tattgacctg 480
aaaaaaaaatcg actctgatac tctgggcctg aatggtttta acgtgaatgg ttctgggacg 540
attaccaaca aagcagcaac tgtcagtgat gttactcgcg caggcggtac attggtgaat 600
ggtgcctatg atataaaaac cactaacaca gcgctgacta caactgatgc ctccgcgaaa 660
ttgaatgatg gtgatgttgt tactatcaat aatggtaagg atactgccta taaatataat 720
gctgctacag gtgggtttac gacggatgtc tccatctccg gggatcctac cgctgctgac 780
gctactgcta ataaaactgc ccgtgatgca cttgcggcgt ctttacatgc tgagccgggt 840
aaaactgtta atggttcttg gactacgaat gatggtacgg taaaatttga taccgatgcc 900
gatggtaaga tttctatttg tgggtgtgct gcttatgtag atgcagcagg caacctgacc 960
actaacgcag caggatatgac gactcaagca acaactaccg atttggttac tgctgctgca 1020
tctgctactg gtaaggggtg atccctgacc tttggtgaca cgacgtataa aattggtcag 1080
ggtacggctg gggttgatcc tgatgacgct tcagatgatg tactgggcac catttcttac 1140
tctaaatcag taagcaagga tgttgttctt gctgatacta aagcaactgg taacacgaca 1200
acagttgatt tcaactccgg tatcatgact tcaaagggtta gtttcgatgc aggtacatca 1260
actgatacat tcaaagatgc agatggtgct atcaccaaaa ctaaagaata caccacttct 1320
tatgctgtaa ataaagatac tgggtgaagtt accgttgctg attatgctgc ggtagatagc 1380
gccgataagg ctgttgatga tactaaatat aaaccgacta tcggcgcgac agttaacctg 1440
aattctgcag gtaaatgac cactgatacc accagtgcag gcacagcaac caaagatcct 1500
ctggctgccc tggacgctgc tatcagctcc atcgacaaat tccgttcato cctgggtgct 1560
atccagaacc gtctggattc cgcagtcacc aacctgaaca acaccactac caacctgtcc 1620
gaagcgcagt cccgtattca ggacgccgac tatgcgaccg aagtgtccaa catgtcgaaa 1680
gcgcagatta tccagcaggc cggtaactcc gtgctggcaa aagccaacca ggtaccgcag 1740
caggttctgt ctctgctaca ggggttaa 1767

- 15 -

<210> 23

<211> 1383

<212> DNA

<213> Escherichia coli

<400> 23

```

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tttcttctgg tctgcgtatt 60
aacagcgcta aagatgacgc tgcgggccag gcgattgcta accgcttcac ttctaacatc 120
aaaggctctga ctcaggccgc acgtaacgcc aacgacggtta tttctctggc gcagaccact 180
gaaggcgcgc tgtctgagat taacaacaac ttgcagcgtg tgcgtgagtt gactgtacag 240
gcgacgaccg ggactaactc tgattctgac ctgtcttcta tccaggatga aatcaaatcc 300
cgtttaagcg aaattgaccg tgtatctggt cagactcagt ttaacggcgt gaacgtactg 360
gctaagaatg acaccctgtc tattcaggta ggtgcaaag acggtcagac tatcaatatt 420
gacctgcagc aaatcgattc tcatacactg ggtctggatg gtttcagcgt taaaaataat 480
gatgcagtga aaaccagtgc tgccgtgaat actcttgggg ggggggcagg ttctgttgct 540
gtcgacttcg caacaaccag ttgactgct atcactggtc tcggtagcgg tgctatcagc 600
gaaattgcta aagacgataa tgggtgattac tacgcgcag tcacagggac tacgggtaat 660
actgctgatg gttactatgc tgcgatatc gacaaggcta ccggtgaggt cgctctgaaa 720
gatggtaacg tagatacacc gacaggtagc ccaacgacga caagcacata tgacttcaca 780
gacgctggtc aaaccgtttc ctttggcact gatgctgcaa cagccggtat cagcactggg 840
gcttctctcg ttaaacttca ggatgagaaa ggcaatgata ctgctactta tgcaatcaaa 900
gcacaagatg gcagcctgta tgccgccaac gttgatgagg ctaccggtaa agtcactgtc 960
aaaaccgcca gctatactga tgctgacggc aaagcagtga ccgatgccgc tgtaaaactg 1020
gggtggtgaca atggcacaac cgaaattggt gtcgatgctg cgtcaggtaa aacttacgat 1080
gctggtgcac tgcaaaacgt tgatctctcc agtgcaacca acacggtaac cgcaatcccg 1140
aacggtaaaa ccacgtctcc gctggctgcc cttgacgacg caatcagcca gatcgacaaa 1200
ttccgctcct cctcgggtgc ggtgcagaac cgtctggatt ccgcggtcac caacctgaac 1260
aacaccacta ccaacctgtc tgaagcgag tcccgatttc aggacgctga ctatgcgacc 1320
gaagtatcca acatgtcgaa agcgcagatc atccagcagg caggtaactc cgtgctgtcc 1380
aaa

```

<210> 24

<211> 1197

<212> DNA

<213> Escherichia coli

<400> 24

```

gcgctgtcga cttctatcga gcgcctctct tctgggtctgc gcattaacag cgctaaagat 60
gacgctgcgg gccaaagcat tgctaaccgc ttcacttcta acatcaaagg tctgactcag 120
gccgcacgta acgccaacga cggattttct ctggcgacga cactgaagg cgcactgtct 180
gaaatcaaca acaacttgca gcgtgttcgt gaactgaccg ttcaggccac taccgggtact 240
aactctgatt ctgacctgtc ttcaatacag gacgaaatca aatcccgctc cgatgaaatt 300
gaccgcgtat ccggtcagac tcagttcaac ggcgttaatg ttctttccaa agatggttca 360
atgaaaattc aggttggtgc gaatgatggt caaactatct ccatcgatct gaagaaaatt 420
gatttttcaa ctttggggct gaatggcttc tcagtttcta aaaactctct taatgtcagc 480
aatgctatca catctatccc gcaagccgct agcaatgaac ctggtgatgt taacttcggg 540
gatactgatg agtctgcagc aatcgcagcc aaattggggg ttccgatac gtcaagccctg 600

```

tcgctgcaca acatccttga taaagatggt aaggcaacag ctgattatgt tgttcagtca 660
ggtaaagact tctatgctgc ttctgttaat gccgcttcag gtaaagtaac cttaaacc 720
attgatgtta cttatgatga ttatgcgaac ggtgttgacg atgccaagca aacaggtcag 780
ctgatcaaag tttcagcaga taaagacggc gcagctcaag gttttgtcac acttcaaggc 840
aaaaactatt ctgctggtga tgcggcagac attcttaaga atggagcaac agctcttaag 900
ttaactgatc tgaatttaag tgatgttact gatactaag gtaaggtaac cacaactgcg 960
actgagcaat ttgaagggtgc ttcaactgag gatccgctgg cgcttctgga taaagctatt 1020
gcacagtcg acaaattccg gtcttctcta ggtgccgtgc agaaccgtct cgattccgct 1080
atcaccaacc tgaacaacac caccaccaac ctgtctgaag cgcagtcctc tattcaggac 1140
gccgactatg cgaccgaagt gtccaacatg tcgaaagcgc agatcatcca gcaggca 1197

<210> 25

<211> 1674

<212> DNA

<213> Escherichia coli

<400> 25

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattctgt aactgacggt tcaggccact 300
acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctg 360
gacgaaattg acccgctatc tggctcagacc cagttcaacg gcgtgaacgt gctgtctaaa 420
gatggctcga tgaaaattca ggtcggcgcg aacgatggcg aaacgattac tattgatctg 480
aagaaaattg actctgatac gctaaatctg gctgggttta acgtgaatgg tgctggctct 540
gttgataatg ccaaggcgac tggcaaagat cttactgatg ctgggtttac ggcaagcgca 600
gctgatgcta atggcaaaat cacttatacc aaagacaccg ttactaaatt cgacaaagcg 660
acagcggctg atgtattggg caaagcggct gctggcgata gcattaccta tgcgggcact 720
gatactggct taggagtcgc tgctgatgcc tcgacttaca cctacaatgc agccaataag 780
tcttacactt ttgatgctac tgggtgttgc aaggcggatg ctggaacggc actgaaagg 840
tacttaggcg catctaaccac cggtaaaatt aatatcggtg gtaccgagca agaagttaac 900
attgccaaag atggctccat caccgatacc aatggcgatg cgctgtatct cgatagtacc 960
ggcaacttaa ccaaaaatac cgcaatttg gggcgtgctg ataaagcaac tgtagataaa 1020
ctgtttgctg gtgctcagga tgcaacgac accttcgata gcggcatgac agctaaattc 1080
gatcaaaactg ctggtaccgt tgatttcaaa ggcgcgtcta tttctgctga tgcaatggca 1140
tcaaccttaa ataattggtc ctatacagcc aacgtagggt gtaaggctta tgccgtaacc 1200
gctggcgcag ttcagacagg tggcgcagat gtgtataaag ataccactgg cgcactgacg 1260
actgaagatg acgaaaccgt taccgcgacc tactacggtt ttgctgatgg taaagtttct 1320
gacggtgaag gttctactgt ctataaagct gctgatgggt ccatcactaa agatgcgact 1380
accaagtctg aagcaaccac tgacctctg aaagcccttg acgacgcaat cagccagatc 1440
gacaaattcc gctcctccct cgggtccggt caaaaccgtc tggattccgc cgtcaccaac 1500
ctgaacaaca ccactacca cctgtctgaa gcgcagtcct gtattcagga cgcgcactat 1560
gcgaccgaag tgtccaacat gtcgaaagcg cagatcatc agcaggccgg taactccgtg 1620
ctggcaaaag ccaaccaggt accgcagcag gttctgtctc tgctgcaggg ttaa 1674

<210> 26

<211> 1365

- 17 -

<212> DNA

<213> Escherichia coli

<400> 26

aacaaatctc agtcttctct tagctctgct attgagcgtc tctcttctgg cctgcgtatt 60
aacagtgcta aagatgacgc agcaggtcag gcgattgcta accgttttac ggcaaattatt 120
aaaggctctga ctcaggcttc ccgtaacgcg aatgatggta tttctgttgc gcagactact 180
gaagggtgcgc tgaatgaaat taacaacaac ctgcagcgtg tacgtgaact gactgttcag 240
gcaactaacg gtactaactc tgacagcgat ctttcttcta ttcaggcaga aattactcaa 300
cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360
gccgaaaata atgaaatgaa aattcagggtt ggtgctaata atggggaaac catcactatc 420
aatctggcaa aaattgatgc gaaaactctc ggccctggacg gctttaatat cgatggcgcg 480
cagaaagcaa ctggcagtgta cctgatttct aaatttaaag cgacaggtag tgataattat 540
caaattaacg gtactgataa ctatactgtt aatgtagata gtggagcagt tcaaatgag 600
gatggtgacg caatttttgt tagcgctacc gatggttctc tgactactaa gagtgataca 660
aaagtcggtg gtacaggtat tgatgcgact gggcttgcaa aagccgcagt ttcttttagct 720
aaagatgctt caattaaata ccaagggtatt actttcacca acaaaggcac tgatgcattt 780
gatggcagtg gtaacggcac tctaaccgct aatattgatg gcaaagatgt aacctttact 840
attgatgcga caggggaagga cgcaacatta aaaacgtctg atcctgttta caaaaatagt 900
gcaggtcagt tcactacaac taagggttgaa aacaaagccg ctacagcatc ggatctggac 960
ttaaataacg ctaaaaaagt gggtagttct ttagttgtaa atggcgctga ttatgaagt 1020
agcgtgatg gtaagacagt aactgggctt ggcaaaacta tgtatctgag caaatcagaa 1080
gggtgtagcc cgattctggt aaaagaagat gcagcaaaat cgttgcaatc tactaccaac 1140
ccgctcgaaa ccacgcgaaa ggcattggct aaagttgaca atctgcgttc tgacctcgg 1200
gcagtacaaa accgtttcga ctctgctatc accaaccttg gcaacaccgt aaacaacctg 1260
tcttctgccc gtagccgtat cgaagatgct gactacgcga ccgaagtgtc taacatgtct 1320
cgtgcgcaga tcctgcaaca agcgggtacc tctgttctgg cgcag 1365

<210> 27

<211> 1740

<212> DNA

<213> Escherichia coli

<400> 27

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcgcgagg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgat ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggg tcaggcttct 300
accgggacta actccgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc tggccagacc cagttcaacg gcgtgaacgt actggcgaaa 420
gacggttcaa tgaaaattca ggttggtgag aatgacggcc agactatcad gattgatctg 480
aagaaaattg actctgatac gctggggctg agtgggttta atgtgaatgg tagcggggct 540
gtggctaata ctgcagcgac taaatctgat ttggcagcag ctcaactctt ggctccagg 600
actgctgatg ctaatggtac agttacctat actgttggcg caggcctgaa aacatctaca 660
gctgcagatg taattgagag ttggctaata aacgcaaaag ttaatgccac aattgcaaat 720
ggttttggat cgccaacagc tacagattat acatacaaca gcgctacagg cgattttaca 780
tatagtgcga ctattgcagc tgggtacaaat tctggtgata gtaacagtgc tcagttacaa 840

tccttcctga caccaaaagc gggcgatact gctaacttaa acgttaaaat tggttctacg 900
tcaattgacg ttgtattggc tagcgacggg aaaattaccg cgaaagatgg ttcagaacta 960
tttattgacg tagatggtaa cctcactcaa aacaatgctg ggactgtcaa agcagccact 1020
cttgatgcac tgactaaaaa ctggcataca acaggcacac cgagtgccgt atctacggta 1080
attacaactg aagatgaaac aaccttcact ctggctggcg gtactgatgc tactacttct 1140
ggtgcaatca ctgtagcaaa tgcaagaatg agtgctgagt ctcttcaatc ggcaactaag 1200
tccacaggat tcacagttga tgttgagct actggtacca ggcgaggcga tattaagtt 1260
gatagtaag gtatagtaca acaacacaca ggtacagggt ttgaagacgc ttacaccaa 1320
gctgatgggt cactgactac cgataatata accaatctgt ttttgcaaaa agacggaact 1380
gtgaccaatg gttcaggtaa agcagtcctat gtttcagcgg atggtaattt tactactgac 1440
gctgaaacta aagctgcaac caccgccgat ccactgaaag ctctggacga agcgatcagc 1500
tccatcgaca aattccgttc ttccctcggt gcggtgcaaa accgtctgga ttccgcagtc 1560
accaacctga acaacaccac tactaacctg tctgaagcgc agtcccgtat tcaggacgct 1620
gactatgcga ccgaagtgtc caatatgtcg aaagcgcaga tcatccagca ggccggtaac 1680
tccgtgctgg caaaagctaa ccaggtagcg cagcagggtc tgtctctgct gcagggttaa 1740

<210> 28

<211> 1233

<212> DNA

<213> Escherichia coli

<400> 28

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgcatt 60
aacagcgcta aagatgacgc tgcggggccag gcgattgcta accgcttcac ttctaacatc 120
aaagggtctga ctcaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180
gaaggcgcac tgtctgaaat caacaacaac ttgcagcggt ttctgtgagct gaccgttcag 240
gccactaccg gtactaaact tgattctgac ctgtcttcaa tccaggacga aatcaaatcc 300
cgtctcgatg aaattgaccg cgtatccggt cagactcagt tcaacggcgt gaacgtactg 360
gcaaaagata acaccatgaa gattcagggt ggtgcgaacg atggtcagac tatatccatc 420
gacctgcaaa aaatcgactc ttctactctt ggtttgaacg gtttctccgt ttctaaaaat 480
gctctcgaaa ctagcgaagc gatcactcag ttgccgaacg gtgcgaatgc accaatcgct 540
gtgaagatgg atgcgtctgt tctgaccgat cttaacatta ctgatgcttc cgctgtttcg 600
ctgcacaacg taactaaagg tgggtgtcgca acgtctactt atgttggttca gtatggcgat 660
aagagctatg cagcatctgt tgatgcggga ggtacagtaa aactgaataa agccgacgta 720
acataatacg acgcagcaaa tgggtgttacg aatgccaccc agattggtag tctggttcag 780
gttggtgctg atgcaaaaca tgatgcagtt ggttttggta ccgtgcaggg gaaaaactat 840
gttgctaatt actcattagt caatgctaatt ggcgctgctg gcgctgcagc aactagagtt 900
acaattgatg gtgatggtag ccttgagct aaccaggcta aaattgaact tagccaaaat 960
ggtgctactg ctgcaacatc agagtctgct ggtgcttcaa ccaacgatcc actgactctg 1020
ctggacaaag ctatcgcatc tgttgataaa ttccggttct ctttgggggc ggtacagaac 1080
cgtctgagct ccgctgtaac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1140
tcccgatttc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1200
atccagcagg caggtaaact cgtgctgtcc aaa 1233

<210> 29

<211> 1713

<212> DNA

<213> Escherichia coli

- 19 -

<400> 29

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggt tcaggcgacg 300
accggaacta actccacctc tgacctggac tccattcagg acgaaatcaa atcccgtctt 360
gatgaaattg accgcgtatc cggccaaacc cagttcaacg gcgtgaacgt actgtcaaaa 420
gatggctcga tgaaaattca ggtcggcgca aatgatggtg aaaccatcac gattgatctg 480
aaaaagatcg actcttctac attgaagctg accagcttca atgttaacgg taaaggcgct 540
gttgataatg ctaaagccac tgaagcagat ctgaccgctg cgggcttctc ccaaggtgca 600
gtcgtcagtg gcaacagcac ctggactaaa tctactgtta ctacctttaa tgcagcaaca 660
gctaccgacg tgctggcaag cgtttagcggc ggcagcacta ttagcgggta taccgggtaca 720
aacaatggat taggcgtagc ggcttctact gcatatacct acaacgcaac cagcaagtct 780
tattcatttg acgcaaccgc acttaccat ggcgatggta ctggggccac cactaaagtt 840
gctgatgtgc tgaaagccta tgcagcaaac ggtgataata cggctcagat ctccatcggc 900
ggaagcgctc aggacgttaa aattgccagc gatggcacc tgactgacgt caatggtgat 960
gctttatata ttggttctga cggcaacctg actaaaaacc aggcggcggt tccagatgcg 1020
gcaacgttgg acggtatttt caacgggtgcg aatggtaatg cagcagttga tgcgaagatt 1080
acattcggca gcggcatgac cgttgatttc acccaggcta gcaaaaaagt ggatattaag 1140
ggcgcaacgg tatccgccga agatatggac actgcgttaa ctgggcaggc ttataccgta 1200
gctaacggcg cacagtcttt tgacgttgcc gctggtgggg cagtaaccgc tactacaggt 1260
ggcgctaccg taaatattgg tgctgatggt gaactgacga ctgcgaccaa caagactgtc 1320
acagaaactt atcacgaatt tgctaacggc aatattctgg atgatgacgg cgcggtctctg 1380
tacaaagcgg ctgacggttc tctgaccact gaagctactg gtaaattcga agtgaccacg 1440
gatccgctga aagcgctgga cgatgctatc gcatccgtag acaaattccg ctctctccctc 1500
ggtgcggtgc agaaccgtct ggattccgca gtcaccaacc tgaacaacac cactaccaac 1560
ctgtctgaag cgcagtcctc cattcaggac gccgactatg cgaccgaagt gtccaatatg 1620
tcgaaagcgc agatcatcca gcaggccggt aactccgtgc tggcaaaagc caaccaggta 1680
ccgcagcagg ttctgtctct gctgcagggt taa 1713

<210> 30

<211> 1668

<212> DNA

<213> Escherichia coli

<400> 30

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gctaaggatg acgcccgagg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggcttct 300
accgggacta actccgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc tggccagacc cagttcaacg gcgtgaacgt actggcgaaa 420
gacggttcaa tgaaaattca ggttggtgcg aatgacggcc agactatcac tattgatctg 480
aagaaaattg actcagatac gctggggctg agtgggttta atgtgaatgg tggcggggct 540
gttgctaata ctgcagcgac taaagatgat ttggctgctg catcagtttc agctgcggta 600

ggtaatgaat acactgtctc tgctggcctg tcgaaatcaa ctgctgctga tgttattgct 660
agtctcacag atgggtgcgac agtaactgcg gctgggtgtaa gcaatgggtt tgctgcaggg 720
gcaactggag atgcttataa attcaatcaa gcaaacaaca cttttactta caataccacc 780
tcaacagcgg cagaactcca atcttacctc acgcctaagg cgggggatac cgcaacttcc 840
tccgttgaaa ttgggtggcac caagcaggat gttgttctgg ctagtgatgg caaaatcaca 900
gcaaaagacg ggtctaaact ttatattgac accacagggg atttaaccca aaacgggtga 960
ggtacttttag aagaagctac cctcaatggc ttagctttca accactctgg tccagccgct 1020
gctgtacaat ctactattac tactgcggat ggaacttcaa tagttctagc aggttctggc 1080
gacttttgaa caacaaaaac tgctggggct attaattgca caggagcagt gatcagtgtc 1140
gatgcacttc ttccgcgcag taaagcgact ggggttactt ctggcactta taccgtaggt 1200
acagatggag ttgttaaact tgggtgcaat gacgtttata acaaagctga cgggacggga 1260
ttaactactg acaataccac aaaatattat ttacaagatg acgggtctgt aactaatggt 1320
tctggtaaag ctgtgtatgc tgatgcaaca ggaaaactaa ctactgacgc tgaaaactaa 1380
gccgaaacca ccgccgatcc cctgaaagct ctggacgaag cgatcagctc catcgacaaa 1440
ttccgttctt ccctcggtgc ggtgcaaaac cgtctggatt ccgcggtcac caacctgaac 1500
aacaccacta ccaacctgtc cgaagcgagc tcccgtattc aggacgccga ctatgcgacc 1560
gaagtgtcca acatgtcgaa agcgcagatc atccagcagg ccggttaactc cgtgctggca 1620
aaagctaacc aggtaccgca gcaggttctg tctctgctgc agggttaa 1668

<210> 31

<211> 1713

<212> DNA

<213> Escherichia coli

<400> 31

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttccg ttgcgcagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggccact 300
accggtacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360
gatgaaattg accgcgtatc tggctcagacc cagttcaatg gcgtgaatgt gttgtccaaa 420
gacggttcaa tgaaaattca ggtgggcgca aatgatggtg aaaccatcac gattgacctg 480
aaaaaaaaatc actcttctac actgaagctg accagcttca acgtcaacgg taaaggcgct 540
gttgataatg caaaagccac tgaagcagat ctgaccgctg cgggcttctc ccaaagtga 600
gttgctcagt gcaatagcac ctggactaaa tctactgtta ctacctttaa tgcagcaaca 660
gctaccgatg tgctggctag cgttagtggc ggcagcacta ttagcgggta tgctggcaca 720
aacaatgggt taggcgtagc ggcttctact gcatatacct acaacgcaac cagcaagtct 780
tattcatctg acgcaaccgc acttactaat ggtgatggtg ctgcgggctc aactaaagt 840
gctgatgttc tgaaagccta tgcagcaaac ggcgataaca cggctcagat ctccatcggt 900
ggtagcgctc aggaagttaa aattgccagc gatggtaccc tgacggatac taatggcgat 960
gctttatata ttggtgctga cggtaacctg acgaaaaacc aggcggcgcg cccagccgcg 1020
gcaacgttgg acggtatttt caacggtgcg aatgggtcat atgcagttga tgcgaagatt 1080
accttcggca gcggcatgac cgttgacttc acccagggtt gcaacaatgt ggatattaag 1140
ggcgcgacgg tatccgccga agatatgaac actgcgttaa ccggtcaggc ttataccgta 1200
gctaacggcg cacagtctta tgacgttgcc gctgatggtg cagtaactgc tactacaggt 1260
ggagcgaccg taaatattgg tgctgagggt gaactgacga ctgcggccaa caagactgtc 1320
acagaaaact atcacgaatt tgctaaccgc aatattctgg atgatgacg cgcggtctg 1380

tataaagcgg ctgacggctc tctgaccact gaagctacag gtaaattctga agcgaccacg 1440
gatccgctga aagcgctgga cgatgctatc gcatccgtag acaaattccg ttcttccctg 1500
ggtgccgtgc agaaccgtct ggattccgca gtcaccaacc tgaacaacac cactaccaac 1560
ctgtccgaag cgcagtcocg tattcaggac gccgactatg cgaccgaagt gtccaacatg 1620
tcgaaagcgc agattattca gcaggcagggt aactccgtgc tggcaaaagc taaccaggta 1680
ccgcagcagg ttctgtctct gctgcagggt taa 1713

<210> 32

<211> 1188

<212> DNA

<213> Escherichia coli

<400> 32

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgcatt 60
aacagcgcta aagatgacgc tgcggggccag gcgattgcta accgcttcac ttctaacatc 120
aaagggtctga ctcaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180
gaaggcgcac tgtctgaaat caacaacaac ttgcagcgtg tgcgtgagtt gactgttcag 240
gcgacgaccg ggactaactc tgattctgac ctgtcttcta ttcaggacga aatcaaattc 300
cgtctggatg aaattgaccg tgtttccggg cagaccaggt tcaacggcgt gaacgtgctg 360
gctaaaaacg gttctatggc gattcagggt ggcgcgaaat atgggcagac catcaacatc 420
gacctgcaga aaatcgactc ttctactctg ggccctggcg gcttctccgt atctaacaat 480
gcactgaaac tgagcgattc tatcactcag gttgggtgca gtggttcact ggcagatgtg 540
aaactgagct ctggtgcctc ggctctgggt gtagacgcaa gcactctgac tctgcacaac 600
gtacagaccc cagctggcgc agcaacagct aactatgttg tctcttctgg ttctgacaac 660
tactcagtat ctggtgaaga tagctccggt acagttacgc tgaacaccac tgatataggt 720
tataccgata ccgctaattg cgttactacc ggttccatga ctggtaagta cgttaaagtt 780
ggagctgatg cattgggtgc tgctgtaggt tatgtcaccg tacagggaca aaacttcaaa 840
gctgatgctg gcgcgctggg taactccaag aatgctgctg gtagtcagaa tgttacttct 900
gcaattggcg atattgctaa taaagcgaat gctaacaatt acactggaac ctcttctgca 960
gatccactgg ctctgctgga caaagctatc gcactctgtg ataaattccg ttcttctcta 1020
ggggcggtgc agaaccgtct gagctctgct gtaaccaacc tgaacaacac cactaccaac 1080
ctgtccgaag cgcagtcocg tattcaggac gccgactatg cgaccgaagt gtccaacatg 1140
tcgaaagcgc agatcatcca gcaggcgggt aactccgtgc tgtctaaa 1188

<210> 33

<211> 1638

<212> DNA

<213> Escherichia coli

<400> 33

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcggc tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaatgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggt tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgctc 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaaaattca ggttggtgcg aacgacggcc agactatcac tattgatctg 480

aagaaaattg actctgatac gctggggctg agtgggttta acgtaaattg tagcgcagat 540
aaggcaagtg tcgcggcgac agctgacgga atgggttaaag acggatatat caaaggggta 600
acttcatctg acggcagcac tgcatatact aaaactacag caaatactgc agcaaaagga 660
tctgatattc ttgcggcgct taagactggc gataaaatta ccgcaacagg tgcaaatagc 720
cttgctgata atgcgacatc gacaacttat acttataatg caaccagcaa taccttctcc 780
tatacggctg acggtgtaaa ccaaacgaat gctgcagcaa atctcatacc tgcagcaggg 840
aaaacgacag ctgcatcagt tactattggg gggacagcac agaattgtaa tattgatgat 900
tcgggcaata ttacttcaag tgatggcgat caactttatc tggattcaac aggtaacctg 960
actaaaaacc aggccggcaa cccgaaaaaa gcaaccgttt ctgggcttct cggaaatacg 1020
gatgcgaaag gtactgctgt taaaacaacc atcaagacag aggctgggtg aacagttaca 1080
gctgaaggta atacaggta tgtaaaaatt gaagggtgta ctgtttcagc atctgcattt 1140
acgggcattg catattccgc caacaccggg ggggaatactt atgctgttgc cgcaataat 1200
actacaaatg gtttctctggc ggggggatgac ttaaccagg atgctcaaac tgtttcaacc 1260
tactactcgc aagccgatgg cacgggtcac aatagcgag gcaaagaaat ctataaagac 1320
gctgatgggtg tctacagcac agagaataaa acatcgaaga cgtccgatcc attggctgcg 1380
cttgacgacg caatcagctc catcgacaaa ttccgttcat ccttgggtgc tatccagaac 1440
cgtctggatt ccgcggtcac caacctgaac aacaccacta ccaacctgtc cgaagcgag 1500
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1560
atccagcagg ccggtaactc cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1620
tctctgctgc agggctaa 1638

<210> 34

<211> 2145

<212> DNA

<213> Escherichia coli

<400> 34

aacaaatctc agtcttctct gagctccgcc attgaacgtc tctcttctgg cctgcgtatt 60
aacagtgtga aagatgacgc agcaggtcag gcgattgcta accgttttac agcaaatatt 120
aaagggtctga ctcaggcttc ccgtaacgcg aatgatggta tttctgttgc gcagaccact 180
gaagggtgcg tgaatgaaat taacaacaac ctgcagcgtg tacgtgaact gactgttcag 240
gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggctga aattactcaa 300
cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360
gctgaaaata atgaaatgaa aattcagggtt ggtgctaata atgggtgaaac catcactatc 420
aatctggcaa aaattgatgc gaaaactctc ggcttgacg gttttaatat cgatggcgcg 480
cagaaagcaa ctggcagtga cctgatttct aaatttaaag cgacagggtac tgataactat 540
gatgttggcg gtgatgctta tactgttaac gtagatagcg gagctgggta atgactcaa 600
cttattgata gtgttttatg ttcagataat gccgatgac tttgtcatgc agctccaccg 660
attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc agattcagggt 720
tatgccgctc aattcgcgtc gtatatcgct tgctgattac gtgcagcttt cccttcaggc 780
gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag ggtgacagca 840
ggctcataag acgcccagc gtcgccatag tgcgttcacc gaatacgtgc gcaacaaccg 900
tcttccggag cctgtcatac gcgtaaaaca gccagcgtg gcgcgattta gcccgcacat 960
agtccactg ttctgctcatt tccgcgcaga cgatgacgtc actgcccggc tgtatgcgcg 1020
aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga ggccaacgcc 1080
cataatgcgg gcagttgccc ggcatccaac gccattcatg gccatatcaa tgattttctg 1140
gtgcgtaccg ggttgagaag cggtgtaagt gaactgcagt tgccatgttt tacggcagtg 1200
agagcagaga tagcgtgat gtcggcggtt gcttttgccg ttacgcacca ccccgctcgt 1260

agctgaacag gagggacagc tgatagaaac agaagccact ggagcacctc aaaaacacca 1320
 tcatacacta aatcagtaag ttggcagcat taccgcggag ctgttaaaga tactacaggg 1380
 aatgatattt ttgttagtgc agcagatggg tcaactgaaa ctaaactctga cacaacata 1440
 gctggtacag ggattgatgc tacagcactc gcagcagcgg ctaagaataa agcacagaat 1500
 gataaattca cgtttaatgg agttgaattc acaacaacaa ctgcagcggg tggcaatggg 1560
 aatggtgtat attctgcaga aattgatggg aagtcagtga catttactgt gacagatgct 1620
 gacaaaaaag cttctttgat tacgagtggg acagtttaca aaaatagcgc tggcctttat 1680
 acgacaacca aagttgataa caaggctgcc acactttccg atcttgatct caatgcagct 1740
 aagaaaacag gaagcacgtt agttgttaac ggtgcaactt acgatgttag tgcagatggg 1800
 aaaacgataa cggagactgc ttctggtaac aataaagtca tgtatctgag caaatcagaa 1860
 ggtggtagcc cgattctggg aaacgaagat gcagcaaaat cgttgcaatc taccaccaac 1920
 ccgctcgaaa ctatcgacaa agcattgggt aaagttgaca atctgcgttc tgacctcggg 1980
 gcagtacaaa accgttttga ctctgctatc accaaccttg gcaacaccgt aaacaacctg 2040
 tcttctgccc gtagccgtat cgaagatgct gactacgcga ccgaagtgtc taacatgtct 2100
 cgtgcgcaga tcctgcaaca agcgggtacc tctgttctgg cgcag 2145

<210> 35

<211> 1587

<212> DNA

<213> Escherichia coli

<400> 35

aacaagaacc agtctgcgct gtcgagttct atcgagcgtc tgtcttctgg cttgctgatt 60
 aacagcgcga aggatgacgc cgcaggctcag gcgattgcta accgttttac ttctaact 120
 aaaggcctga ctcaggctgc acgtaacgcc aacgacggta tttctgttgc gcagaccacc 180
 gaaggcgcgc tgtccgaaat caacaacaac ttacagcgtg tgcgtgaact gaccgttcag 240
 gcaaccaccg gtaccaactc ccagtctgac ctggactcta tccaggacga aattaaatcc 300
 cgtctggacg aaattgaccg cgtatccggg cagacccagt tcaacggcgt gaacgtactg 360
 gcaaaagacg gttccatgaa aattcagggt ggccgcgaac atggccagac catcactatc 420
 gacctgaaga agattgactc ttctacgctg aaactgactg gttttaacgt gaatggcaaa 480
 gcagcgggtg ataattgtaa agcgacggat gcaaatctga ctaccgccgg ttttacacaa 540
 ggcgttgtgg attcaaatgg taatagtact tggactaaat caactacgac taatttcgat 600
 ggcgcaactg cagtaaacgt actagcagca gttaaagatg gcagcacaat caattacacc 660
 ggtactggta atggttttag gattgctgca acaagtgtt atacatatca cgatagcact 720
 aaatcctata cttttgattc tacgggggct gcagtagctg gtgccgcgtc cagcctgcaa 780
 ggtacttttg gtacagatac gaatactgca aaaatcacca tcgatgggtc tgctcaagaa 840
 gtaaaccatc ctaaaagtgg gaaaattact gatactgatg gtaaagcttt atatatcgat 900
 tccactggta atttgactaa gaacggctct gatactttaa ctcaggcaac attgaatgat 960
 gtccttactg gtgctaattc agttgatgat acaaggattg acttcgatag cggcatgtct 1020
 gtcacccttg ataaagtga cagcactgta gatatcactg gcgcacttat ttcagccgct 1080
 gcaatgacta atgagttgac aggtaaggcc tataccgtag taaatggtgc agaactttac 1140
 gctgtagcta ctaataacac agtaaaaacg actgctgatg ctaaaaatgt ttatgttgat 1200
 gctagtggta aattaactac tgatgacaaa gccactgtta cagaaactta tcatgaattt 1260
 gcgaatggca atatctatga tgataaaggc gctgctgttt atgcggcggc ggatggttct 1320
 ctgactacag aaactacaag taaatcagaa gctacagcta acccgctggc cgctctggac 1380
 gacgcaatca gccagatcga caaattccgt tcattccctgg gtgctatcca gaaccgtctg 1440
 gattccgcag tcaccaacct gaacaacacc actaccaatc tgtctgaagc gcagtcctcg 1500
 attcaggacg ccgactatgc gaccgaagtg tccaatatgt cgaaagcgca gatcatccag 1560

caggcaggca actccgtgct ggcaaaa

1587

<210> 36

<211> 1245

<212> DNA

<213> Escherichia coli

<400> 36

aacaaaaacc agtctgcgct gtgcacttct atcgagcgcc tctcttctgg tctgcgcatt 60
aacagcgcta aagatgacgc tgcggggccag gcgattgcta accgcttcac ttctaaccatc 120
aaaggtctga ctcaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180
gaaggcgcac tgtctgaaat caacaacaac ttgcagcgctg ttcgtgaact gaccgttcag 240
gccactaccg gtactaactc tgattctgac ctgtcttcaa tccaggacga aatcaaattcc 300
cgtctcgatg aaattgaccg cgtatccggt cagactcagt tcaacggcgt gaacgtactg 360
gcaaaagatg gctcgatgaa aattcagggtc ggtgcaaattg atggtcagac aatcagcatt 420
gatttgcaga agattgattc ttctacttta gggttaaattg gtttttctgt ttccaaaaat 480
gcagtatctg ttggtgatgc tattactcaa ttgcctggcg agacggcagc cgatgcacca 540
gtaaccatca agtttgatga ttcagtaaaa actgatttaa aactgaccga tgcttcaggg 600
ttaagtctgc ataacctcaa agatgaaaat ggtaatttaa ctaaccagta tgtgttacag 660
aatggcgga aatcttacgc tgctacagtc gctgccaatg gtaatgttac gctgaacaaa 720
gcaaatgtaa cctacagcga tgcgcgaaac ggtattgata ccgcaacgca gtcaggccag 780
ttagttcagg ttggtgcaga ttctaccggt acgccaaaag cattcgtgtc tgtccaagg 840
aaaagctttg gcattgatga cgcgccttg agaataaca ctggtgatgc taccgctact 900
caaccgggaa catctgggac aacagttgtc gcagcgtaa ttcactctgag tacgggcaaa 960
aactctgtag acgtgatgt aacggcttcc actgaattca cagggtgctc aaccaacgat 1020
ccactgactc tgctggacaa agctatcgca tctgttgata aattccgttc ttctttgggg 1080
gcggtacaga accgtctgag ctccgtgta accaacctga acaacaccac caccaacctg 1140
tctgaagcgc agtcccgtat tcaggacgcc gactatgcga ccgaagtgtc caacatgtcg 1200
aaagcgcaga ttatccagca ggcaggtaac tccgtgctgt ccaaa 1245

<210> 37

<211> 1185

<212> DNA

<213> Escherichia coli

<400> 37

aacaaaaacc agtctgcgct gtgcacttct atcgagcgcc tctcttctgg tctgcgcatt 60
aacagcgcta aagatgacgc tgcggggccag gcgattgcta accgcttcac ttctaaccatc 120
aaaggtctga ctcaggctgc acgtaacgcc aatgacggta tttctctagc acagacagcg 180
gaaggcgcg tgcagagat taacaacaac ttgcagcgctg tgcgtgagtt gaccgtgcag 240
gcaaccactg gtaccaactc tgattccgat ctctcttcta ttcaggatga aattaaatct 300
cgtctggatg aaattgaccg cgtctctggt cagaccaggt ttaacggcgt gaacgtactg 360
gctaaaaacg gttctatggc aattcagggt ggcgcgaacg atggccagac tatctctatc 420
gacctgcaga aaatagactc ttctactctg ggtctgagcg gcttctctgt ttctcagaac 480
tccctgaaac tgagcgattc tatcactacg atcggcaata ctactgctgc atcgaagaac 540
gtggacctga gcgcagtagc aactaaactg ggcgtgaatg caagcaccct gagcctgcac 600
gaagttcagg actctgctgg tgacggtaact ggtaccttcg ttgtttcttc tggcagcgac 660
aactatgctg tgtctgtaga cgcggcctct ggtgcagtta acctgaacac cactgacgtc 720

acctatgatg acgctactaa tgggtgttact ggcgcgactc agaacgggtca gctgatcaaa 780
 gtaactttctg acgccaacgg tgacgtctgtt gggttacgtaa ccattcaggg taaaaactat 840
 caggctgggtg cgaccgggtg tgacgttctg gcgaacagcg gtgttgccagc tccaactaca 900
 gctgttgata ccggtactct gcaactgagc ggtactgggtg caactactga gctgaaaggt 960
 actgcaactc agaaccactt ggcaactattg gacaaagcta tcgcttctgt tgataaattc 1020
 cgttcttctc tgggtgcggt acagaatcgt ctgagctctg ctgtaaccaa cctgaataac 1080
 accaccacta acctgtctga agcgcagtc cgtattcagg atgccgacta tgcgaccgaa 1140
 gtgtcaaata tgtctaaagc gcagatcgtt cagcaggccg gtaac 1185

<210> 38

<211> 1383

<212> DNA

<213> Escherichia coli

<400> 38

aacaaatctc agtctctctt tagctctgct attgagcgtc tgtcttctgg tctgcgtatt 60
 aacagcgcaa aagacgatgc agcaggtcag gcgattgcta accgttttac ggcaaatatt 120
 aaaggtctga ccaggtctc ccgtaacgca aatgatggta tttctgttgc gcagaccact 180
 gaaggtgctg tgaatgaaat taacaacaac ctgcagcgta ttcgtgaact ttctgttcag 240
 gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggtga aattactcaa 300
 cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtccct 360
 gctgaaaata atgaaatgaa aattcaggtt ggtgctaata atggtgaaac catcactatc 420
 aatctggcaa aaattgatgc gaaaactctc ggcttgacg gttttaatat cgatggcgcg 480
 cagaaagcaa caggcagtgta cctgatttct aaatttaaag cgacaggtac tgataattat 540
 gatgttggcg gtaaaactta taccgtgaat gtggagagcg gcgcgggttaaat gatgatgct 600
 aataaagatg tttttgtaag cgcagctgat ggatcgctga cgaccagtag tgatactaaa 660
 gtatccggtg aaagtattga tgcaacagaa ctacgcaaac ttgcaataaa attagctgac 720
 aaaggctcca ttgaatacaa gggcattaca tttactaaca aactggcgcg agagcttgat 780
 gctaattgga aaggtgtttt gaccgcaaat attgatggtc aagatgttca atttactatt 840
 gacagtaatg caccacggg tgccggcgca acaataacta cagacacagc tgtttacaaa 900
 aacagtgcgg gccagttcac cactacaaaa gtggaaaata aagccgcaac actctctgat 960
 ctggatctta atgcagccaa gaaaacaggt agcacttttag ttgtaaattg cgccacctac 1020
 aatgtcagcg cagatggtta aacggtaact gatactactc ctggtgcccc taaagtgatg 1080
 tatctgagca aatcagaagg tggtagcccg attctggtta acgaagatgc agcaaatcg 1140
 ttgcaatcta ccaccaaccc gctcgaaact atcgacaagg cattggctaa agttgacaat 1200
 ctgcgttctg acctcggtgc agtacaaaac cgtttcgact ctgccatcac caacctggcg 1260
 aacaccgtaa acaacctgtc ttctgcccgt agcgcgtatc aagatgctga ctacgcgacc 1320
 gaagtgtcta acatgtctcg tgcgcagatc ctgcaacaag cgggtacctc tgttctggcg 1380
 cag 1383

<210> 39

<211> 1680

<212> DNA

<213> Escherichia coli

<400> 39

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120

gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt tcacctctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatctctg ttgcacagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggt tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacggttcaa tgaaaattca ggttggtgcg aatgacggcc agactatcac tattgatctg 480
aagaaaattg actctgatac tctgggtttg agtggattta atgtgaatgg caaaggggct 540
gtggctaacg caaaagcgac cgaagcagat ttaacggggg ctgggtttctc tcaaggagcg 600
gtggatacaa acggaaatag tacttggaca aaatcaacca ccaccaatta ctcagctgca 660
acaactgctg acttggtatc gaccattaag gatggctcta ctgttacata tgcagggaca 720
gacaccggat taggggtcgc agcagcagga aattatactt atgatgcgaa cagtaaactct 780
tattccttca atgccaatgg tctgacgggc gcaaataccg caactgcact caaaggttac 840
ttggggacag gtgctaacac cgctaaaatt tctatcgggt gtacagagca ggaagtgaat 900
attgccaaag atggcactat tacagatacg aatggtgatg cgctctatct ggatattacc 960
ggcaacctga ctaagaacta tgcgggttca ccacctgcag caacgctgga taacgtatta 1020
gcttccgcaa ctgtaaattgc cactatcaag tttgatagcg gtatgacggt tgattacact 1080
gcaggtagct gcgcgaatat tacagggtgca tccatttctg cagatgacat ggccgcaaaa 1140
ctgagcggaa aggcgtacac tgttgccaat ggtgctgagt cttatgacgt tgctgcagtt 1200
acgggggctg taacaactac agcaggtaat tcacctgtgt atgccgatgc agacggtaaa 1260
ttaacgacga gtgccagtaa tacggttact cagacttata acgagtttgc taatggtaac 1320
atztatgatg acaaaggctc gtcactgtat aaagctgcag atggctctct gacttctgaa 1380
gctaaaggga aatctgaagc aaccgcgat cccctgaaaag ctctggacga agccatcagc 1440
tccatcgaca aattccgctc ctcctcgggt gccgttcaaa accgtctgga ttctgcggtg 1500
accaacctga acaacaccac taccaacctg tctgaagcgc agtcccgtat tcaggacgcc 1560
gactatgcga ccgaagtgtc caatatgtcg aaagcgcaga tcatccagca ggccggtaac 1620
tccgtgttgg caaaagctaa ccaggtagcg cagcagggtt tgtctctgct gcagggttaa 1680

<210> 40

<211> 1146

<212> DNA

<213> Escherichia coli

<400> 40

gcgctgtcga cttctatcga gcgcctctct tctgggtttgc gcattaacag cgctaaagat 60
gacgctgcgg gccaggcgat tgctaaccgc ttcacttcta acatcaaagg tctgactcag 120
gccgcacgta acgccaacga cggatatctc ctggcgacga cactgaagg cgactgtct 180
gaaatcaaca acaacttgca gcgtgttcgt gaactgaccg ttcaggccac taccggtact 240
aactctgatt ctgacctgtc ttcaatccag gacgaaatca aatcccgtt ggctgaaatc 300
gatcgtgtct ctggtcagac ccagttcaac ggcgtgaacg tgctggctaa aaacggttct 360
ctgaatattc aggttggcgc gaatgatggg cagaccatct ctatcgattt gcagaaaata 420
gactcttctg cccttgggtt aagtggtttt agtgttgccg gtggggcgct aaaattaagc 480
gatacagtga cgcaggtcgg cgatggttca gccgcgccag ttaaagtgga tctggatgca 540
gcagcaacag atattggtac tgctttgggg caaaagggtta atgcaagttc tttaacgttg 600
cacaatatct tagacaaaga tgggtgcggca actgagaact atgttggttag ctatggtagt 660
gataattacg ctgcatctgt tgcagatgac gggactgtaa ctcttaataa aacggatatt 720
acttattcag gcggtgatat taccggcgct accaaagatg atacgttgat taaagttgct 780
gctaattctg acggagaggg cggttggttc gctaccgttc agggtaagaa ttatgaaatt 840
acagatgggtg taaaaaacca gtccactgct gcaccaaccg atattgctca gaccattgat 900

ctggatacgg ctgatgaatt tactggggct tccactgctg atccactggc acttttagac 960
aaagctattg cacaggttga tactttccgc tctccctcg gtgccgttca aaaccgtctg 1020
gattccgcag tcaccaacct gaacaacact actaccaacc tgtctgaagc gcagtcccg 1080
attcaggacg ccgactatgc gaccgaagtg tccaatatgt cgaaagcgca gatcatccag 1140
caggcc 1146

<210> 41

<211> 1506

<212> DNA

<213> Escherichia coli

<400> 41

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt ttactttctaa tattaagggc 180
ctgactcagg ctgcacgtaa cgccaatgac ggtatttctc tggcgcgagc cactgaaggc 240
gcaactgtctg aaatcaacaa caacttgcag cgtgtgcgtg aactgaccgt acaggcgaca 300
accggaacga actccgaatc tgacctgtcc tctatccagg acgaaatcaa atcccgctctg 360
gaagagattg accgcgtatc cggccagact cagttcaacg gcgtgaatgt gctggcaaaa 420
gacggcacca tgaaaattca ggtaggcgcg aacgatggtc agactatctc tatcgatctg 480
aaaaaaatcg actcttcaac cctgggcctg accggttttg atgtttcgac gaaagcgaa 540
atttctacga cagcagtaac gggggcgcca acgaccactt atgctgatag cgccgttgca 600
attgatatcg gaacggatat tagcggatatt gctgctgatg ctgcgttagg aacgatcaat 660
ttcgataata caacaggcaa gtactacgca cagattacca gtgcgggcaa tccgggcctt 720
gatggtgctt atgaaatcca tgttaatgac gcggtatggt ccttcaactgt agcagcgagt 780
gataaacaag cgggtgctgc tccgggtact gctctgacaa gcggtaaagt tcagactgca 840
accaccacgc caggtagcgc tgttgatgac actgcggcta aaactgctct ggctgcagca 900
ggtgctgaca cgagtggcct gaaactggtt caactgtcca acacggattc cgcaggtaaa 960
gtgaccaacg tgggttacgg cctgcagaat gacagcggca ctatctttgc aaccgactac 1020
gatggcacca ctgtgaccac gccgggcca gagactgtga cttacaaaga tgcttccgg 1080
aacagcacca ctgcggctgt cacactgggt ggctctgatg gcaaaaccaa tctgggtacc 1140
gccgctgacg gcaaaacgta cgggtgcgact gcaactgaatg gtgctgatct gtccgatcct 1200
aataacaccg ttaaactctgt tgcagacaac gctaaaccgt tggctgccct ggatgatgca 1260
attgcatggg tgcacaaatt ccgctcctcc ctcgggtgcgg tgcaaaaccg tctggattcc 1320
gcagtcacca acctgaacaa caccactacc aacctgtctg aagcgcagtc ccgtattcag 1380
gacgccgact atgcgaccga agtgtccaac atgtcgaaag cgcagattat ccagcaggca 1440
ggtaactccg tgctgtccaa agctaaccag gttccgcagc aggttctgtc tctgctgcag 1500
ggttaa 1506

<210> 42

<211> 950

<212> DNA

<213> Escherichia coli

<400> 42

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgtatt 60
aacagcgcta aagatgacgc cgcggggccag gcgattgcta accgctttac ttctaacatc 120
aaaggtctga ctcaggccgc acgtaaccgc aacgacggta tttctctggc gcagacggct 180

gaaggcgcgc tgtcagagat taacaacaac ttgcagcgta ttctggaact gaccgttcag 240
gcctctaccg gcacgaactc tgattccgac ctgtcttcta ttcaggacga aatcaaatec 300
cgtcttgatg aaattgaccg tgtatctggt cagacccagt tcaacgggtg gaacgtgctg 360
tcgaaaaacg attcgaatga gattcagatt ggtgccaatg ataaccagac gatcagcatt 420
ggcttgcaac aaatcgacag taccactttg aatctgaaag gatttaccgt gtcggcatg 480
gcggatttca gcgcggcgaa actgacggct gctgatggta cagcaattgc tgcgggatg 540
gtcaaggatg ctgggggtaa acaagtcaat ttactgtctt acactgacac cgcgtctaac 600
agtactaaat atgcggtcgt tgattctgca accggtaaat acatggaagc cactgtagcc 660
attaccggta cggcggcggc ggtaactgtt ggtgcagcgg aagtggcggg agccgctaca 720
gccgatccgt taaaagcact ggatgccgca atcgctaaag tcgacaaatt ccgtcctcc 780
ctcgggtgcc ttcaaaaccg tctggattct gcggtcacca acctgaacaa caccaccacc 840
aacctgtctg aagcgcagtc ccgtattcag gacgccgact atgcgaccga agtgtccaac 900
atgtcgaaag cgcagattat ccagcaggcc ggtaactccg tgctggcaaa 950

<210> 43

<211> 1707

<212> DNA

<213> Escherichia coli

<400> 43

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt ttacctctaa cattaagggt 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg aactgacggg tcaggcttct 300
accgggacta actccgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc cgggtcaaacc cagttcaacg gtgtgaacgt actggcgaaa 420
gacggttoga tgaaaattca ggttggtgcg aatgaaggcc agactatcac gattgatctg 480
aagaaaattg actcagatac gctggggctg aatgggtttca acgttaatgg caaaggcact 540
attgcgaaca aagctgctac agtcagcgat ctgaccgctg ctggtgcaac gggaacagggt 600
ccttatgctg tgaccacaaa caatacagca ctcagcgcta gcgatgcact gtctcgctg 660
aaaaccggag atacagttac tactactggc tcgagtgtg cgtatctatac ttatgatgctg 720
gctaagggga acttcaccac tcaagcaaca gttgcagatg gcgatgttgt taactttgctg 780
aatactctga aaccagcggc tggcactact gcatacagggt tttatactcg tagtactggt 840
gatgtgaagt ttgatgtaga tgctaattggc gatgtgacca tcggtggtaa agccgcgtac 900
ctggacgcca ctggtaacct atctacaaac aaccccgcca ttgcatcttc agcgaaattg 960
tccgatctgt ttgctagcgg tagtacetta gcgacaactg gttctatcca gctgtctggc 1020
acaacttata actttggtgc agcggcaact tctggcgtaa cctacaccaa aactgtaagc 1080
gctgatactg tactgagcac agtgacaggt gctgcaacgg ctaacacagc agttactggt 1140
gcgacaatta agtataatac aggtattcag tctgcaacgg cgtccttcgg tgggtgtaat 1200
actaatggtg ctggtaattc gaatgacacc tatactgatg cagacaaaga gtcaccaca 1260
accgatctt acactatcaa ctacaacgct gataaggata ccggtacagt aactgtagct 1320
tcaaatggcg cagggtgcaac tggtaaattt gcagctactg ttggggcaca ggcttatgtt 1380
aactctacag gcaaaactgac cactgaaacc accagtgcag gcactgcaac caaagatcct 1440
ctggctgccc tggatgaagc tatcagctcc atcgacaaat tccgttcata cctgggtgct 1500
atccagaacc gtctggattc cgcggttacc aacctgaaca acaccactac caacctgtcc 1560
gaagcgcagt cccgtattca ggacgcgcac tatgcgaccg aagtgtccaa catgtcgaag 1620
gcgcagatta tccagcaggc cggtaaactc gtgctggcaa aagccaacca ggtaccgcag 1680

caggttctgt ctctgctgca ggggttaa

1707

<210> 44

<211> 1720

<212> DNA

<213> Escherichia coli

<400> 44

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccggt ttacttctaa tattaaaggc 180
ctgactcagg ctgcacgtaa cgccaatgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtgtgctg aactgaccgt tcaggcgacc 300
accggtacca actcccagtc tgatctggac tctatccagg acgaaatcaa atcccgtctg 360
gacgaaattg accgcgtatc cggtcagact cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcca tgaaaattca ggttggcgcg aatgatggcc agaccatcac tatcgacctg 480
aagaagattg actcttctac gttgaaactg actggtttta acgtgaatgg ttctggttct 540
gtggcgaata ctgcggcgac taaagacgaa ctggctgctg ctgctgcggc ggccgggtaca 600
actcctgctg tcggtactga cggcgtgacc aaatataccg tagacgcagg gcttaacaaa 660
gccacagcag caaacgtgtt tgcaaacctt gcagatgggt ctggtgttga tgctagcatt 720
tccaacggtt ttggtgcagc agcagccaca gactacacct acaataaagc tacaatgat 780
ttcactttca atgccagcat tgctgctggt gctgcggcgc gtgatagtaa cagcgcagct 840
ctgcaatcct tcttgactcc aaaagcaggt gatacagcta acctgagcgt caaaatcggg 900
acgacatctg ttaatgttgt tctggcgagc gatggcaaaa ttacagcgaa agatggctca 960
gctctgtata tcgactcaac gggtaacctg actcagaaca gcgcaggcac tgtaacagca 1020
gcaaccctgg atggactgac caaaaacat gatgcgacag gagctgttgg tgttgatata 1080
acgaccgcag atggcgcaac tatctctctg gcaggctctg ctaacgcggc aacaggtaact 1140
caatcagggt caattacact gaaaaatgtt cgtatcagtg ctgatgctct gcagtctgct 1200
gcgaaaggta ctgttatcaa tgttgataat ggtgctgatg atatttctgt tagtaaaacc 1260
gggtgtcgtt actaccggag gtgcgcctac ttatactgat gctgatggta aattaacgac 1320
aaccaacacc gttgattatt tctgcaaac tgatggcagc gtaaccaatg gttctggtaa 1380
aggggtttac accgatgcag ctggtaaatt cactaccgac gctgcaacca aagccgcaac 1440
caccaccgat ccgctgaaag cccttgatga cgcaatcagc cagatcgata agttccggtt 1500
atccctgggt gctatccaga accgtctgga ttccgcggtt accaacctga acaacaccac 1560
taccaacctg tccgaagcgc agtcccgat tcaggacgcc gactatgcga ccgaagtgtc 1620
caatatgtcg aaagcgcaga tcatccagca ggccggtaac tccgtgttgg caaaagctaa 1680
ccaggtagcg cagcaggttc tgtctctgct gcagggttaa 1720

<210> 45

<211> 14516

<212> DNA

<213> Escherichia coli

<400> 45

gatctgatgg ccgtagggcg ctacgtgctt tctgctgata tctgggctga gttggaaaaa 60
actgctccag gtgcttgggg acgtattcaa ctgactgatg ctattgcaga gttggctaaa 120
aaacagtctg ttgatccat gctgatgacc ggcgacagct acgactgcgg taagaagatg 180
ggctatatgc aggcattcgt taagtatggg ctgcgcaacc ttaaagaagg ggccaagtgc 240

cgtaagagca tcaagaagct actgagttag tagagattta cacgtctttg tgacgataag 300
 ccagaaaaaa tagcggcagt taacatccag gcttctatgc ttaagcaat ggaatgttac 360
 tgccgttttt tatgaaaaat gaccaataat aacaagttaa cctaccaagt ttaatctgct 420
 ttttgttgga tttttcttg tttctggtcg catttggtaa gacaattagc gtgagtttta 480
 gagagttttg cgggatctcg cggaactgct cacatctttg gcatttagtt agtgcactgg 540
 tagctgttaa gccaggggag gtagcttgcc taattaattt ttaacgtata catttattct 600
 tgccgcttat agcaataaaa gtcaatcgga ttaacttct tttccattag gtaaaagagt 660
 gttttagtagt gtcagggaa attggttttg gtagtagtac tttcaaatt atccattttc 720
 cgatttagat ggcagttgat gttactatgc tgcatacata tcaatgtata ttatttactt 780
 ttagaatgtg atatgaaaaa aatagtgatc ataggcaatg tagcgtcaat gatgttaagg 840
 ttcaggaaaag aattaatcat gaatttagtg aggcaagggtg ataatgtata ttgtctagca 900
 aatgattttt cactgaaga tcttaaagta ctttcgtcat ggggcgttaa gggggttaa 960
 ttctctctta actcaaaggg tattaatcct ttaaggata taattgctgt ttatgaacta 1020
 aaaaaattc ttaaggatat ttccccagat attgtatttt catattttgt aaagccagta 1080
 atatttgga ctattgctt aaagttgtca aaagtgccaa ggattgttg aatgattgaa 1140
 ggtctaggta atgccttcac ttattataag ggaagcaga ccacaaaaac taaaatgata 1200
 aagtggatac aaattctttt atataagtta gcattaccga tgcttgatga tttgattcta 1260
 ttaaatcatg atgataaaaa agatttaatc gatcagtata atattaaagc taaggtaaca 1320
 gtgttaggtg ggattggatt ggatcttaat gagttttcat ataaagagcc accgaaagag 1380
 aaaattacct ttatttttat agcaaggta ttaagagaga aaggatatt tgagtttatt 1440
 gaagccgcaa agttcgtaa gacaacttat ccaagttctg aatttgtaat tttaggaggt 1500
 tttgagagta ataactctt ctcattacaa aaaaatgaaa ttgaatcgct aagaaaagaa 1560
 catgatctta tttatcttg tcatgtggaa aatgttcaag attggttaga gaaaagtct 1620
 gtttttggtt tacctacatc atatcgagaa ggcgtacca ggggtgatcca agaagctatg 1680
 gctattggta gacctgtaat aacaactaat gtacctgggt gtagggatat aataatgat 1740
 ggggtcaatg gctttttgat acctcattt gaaattaatt tactggcaga aaaaatgaaa 1800
 tattttattg agaataaaga taaagtactc gaaatggggc ttgctggaag gaagtttgca 1860
 gaaaaaaact ttgatgcttt tgaaaaaaat aatagactag catcaataat aaaatcaaat 1920
 aatgattttt gacttgagca gaaattattt atatttcaat ctgaaaaata aaggctgtta 1980
 ttatgaataa agtggcatta attactggta tcaactggga agatggctcc tatttggcag 2040
 aattattgtt agaaaaaggt tatgaagttc atggtattaa acgccgtgca tcttcattta 2100
 atactgagcg agtggatcac atctatcagg attcacattt agctaactct aaactttttc 2160
 tacactatgg cgatttgaca gatacttcca atctgacctg ttttttaaaa gaagttcaac 2220
 cagatgaagt ttacaatttg ggggcgatga gccatgtagc ggtatcattt gagtcaccag 2280
 aatacactgc tgatgttgat gcgataggaa cattgcgtct tcttgaagct atcaggatat 2340
 tggggctgga aaaaaagaca aaattttatc aggcctcaac ttcagagctt tatggtttgg 2400
 ttcaagaaat tccacaaaaa gagactacgc cattttatcc acgttcgcct tatgctgttg 2460
 caaaattata tgctatttg atcactgtta attatcgtga gtcttatggg atgtttgcct 2520
 gcaatgggat tctctttaac cagcaatcac ctgcgcgtgg cgagaccttt gttactcgta 2580
 aaataacacg cgggatagca aatattgctc aaggctctga taaatgctta tacttgggaa 2640
 atatggattc tctgctgat tggggacatg ctaaggatta tgtcaaaatg caatggatga 2700
 tgctgcagca agaaactcca gaagattttg taattgctac aggaattcaa tattctgtcc 2760
 gtgagtttgt cacaatggcg gcagagcaag taggcataga gtttagcattt gaaggtagg 2820
 gagtaaatga aaaagggtgt gttgtttcgg tcaatggcac tgatgctaaa gctgtaaac 2880
 cgggcgatgt aattatatct gtagatccaa ggtatttttag gcctgcagaa gttgaaacct 2940
 tgcttggcga tctactaat gcgcataaaa aattaggatg gagccctgaa attacattgc 3000
 gtgaaatggt aaaagaaatg gtttccagcg atttagcaat agcgaaaaag aacgtcttgc 3060
 tgaaagctaa taacattgac actaatattc cgcaagaata aaaaagataa tacattaat 3120

aattaaaaat ggtgctagat ttattagtag cattatTTTT ttttgggtga ctaatgttta 3180
ttacatcaga taaatttaga gaaattatca agttagttcc attagtagtca attgatctgc 3240
taattgaaaa cgagaatggt gaattattat ttggtcttag gaataatcga ccggccaaaa 3300
attatTTTT tgttccaggt ggtaggattc gcaaaaatga atctattaaa aatgctttta 3360
aaagaatata atctatggaa ttaggttaaag agtatggtat ttcaggaagt gtttttaatg 3420
gtgtatggga acatttctat gatgatggtt tttttctga aggcgaggca acacattata 3480
tagtgctttg ttacacactg aaagtcttta aaagtgaatt gaatctccca gatgatcaac 3540
atcgtgaata cctttggcta actaaacacc aaataaatgc taaacaagat gttcataact 3600
attcaaaaaa ttattttttg taatttttat taaaaattaa tatgcgagag aattgtatgt 3660
ctcaatgtct ttaccctgta attattgccc gaggaaccgg aagccgtcta tggccgttgt 3720
ctcgagtatt ataccctaaa caatttttaa atttagttgg ggattctaca atgttgcaaa 3780
caacaattac gcgtttggat ggcacgaat gcgaaaatcc aattgttatc tgcaatgaag 3840
atcaccgatt tattgtagca gagcaattac gacagattgg taagctaacc aagaatatta 3900
tacttgagcc gaaaggccgt aatactgcac ctgccatagc tttagctgct tttatcgctc 3960
agaagaataa tcctaataac gaccctttat tattagtact tgcggcagac cactctataa 4020
ataatgaaaa agcatttctga gagtcaataa taaaagctat gccgtatgca acttctggga 4080
agttagtaac atttgaatt attccggaca cggcataac tgggtatgga tatattaaga 4140
gaagttcttc agctgacct aataaagaat tcccagcata taatgttgcg gagttttag 4200
aaaaaccaga tgttaaaaca gcacaggaat atatttctga tgggaattat tactggaata 4260
gcggaatggt tttatttctc gccagtaaat atcttgatga actacggaaa tttagaccag 4320
atatttatca tagctgtgaa tgtgcaaccg ctacagcaaa tatagatatg gactttgtcc 4380
gaattaacga ggctgagttt attaattgtc ctgaagagtc tatcgattat gctgtgatgg 4440
aaaaaacaag agacgctgta gttcttccga tagatattgg ctggaatgac gtgggttctt 4500
ggctcatcact ttgggatata agccaaaagg attgccatgg taatgtgtgc catggggatg 4560
tgctcaatca tgatggagaa aatagtttta tttactctga gtcaagctctg gttgagacag 4620
tcggagtaag taatttagta attgtccaaa ccaaggatgc tgtactggtt gcggaccgtg 4680
ataaagtcca aaatgttaaa aacatagttg acgatctaaa aaagagaaaa cgtgctgaat 4740
actacatgca tcgtgcagtt tttcgccctt ggggtaaatt cgatgcaata gaccaaggcg 4800
atagatatag agtaaaaaaa ataatagtta aaccaggaga agggtagat ttaaggatgc 4860
atcatcatag ggcagagcat tggattgttg tatccggtac tgctaaagt tccactaggta 4920
gtgaagttaa actattagtt tctaattgag ctatatatat ccctcaggga gcaaaatata 4980
gtcttgagaa tccaggcgta atacctttgc atctaattga agtaagttct ggtgattacc 5040
ttgaatcaga tgatatagtg cgttttactg acagatataa cagtaaaca ttcctaaagc 5100
gagattgata aatatgaata aaataacttg cttcaaagca tatgatatac gtgggcgtct 5160
tggtgctgaa ttgaatgatg aaatagcata tagaattggt cgcgcttatg gtgagttttt 5220
taaacctcaa actgtagttg tgggaggaga tgctcgctta acaagtgaga gtttaagaa 5280
atcactctca aatgggctat gtgatgcagg cgtaaatgtc ttagatcttg gaatgtgtgg 5340
tactgaagag atatattttt ccacttggtt ttttaggaatt gatggtggaa tcgaggtaac 5400
tgcaagccat aatccaattg attataatgg aatgaaatta gtaaccaaag gtgctcgacc 5460
aatcagcagt gacacaggtc tcaaatagat acaacaatta gtagagagta ataattttga 5520
agagctcaac ctagaaaaaa aagggaatat taccaaatat tccaccgag atgcctacat 5580
aatcattttg atgggctatg ctaactctga aaaaataaaa aaaatcaaaa tagttgtgaa 5640
ttctgggaat ggtgcagctg gtctgttat tgatgctatt gaggaatgct ttttacggaa 5700
caatattccg attcagtttg taaaaataaa taatacacc gatggtaatt ttccacatgg 5760
tatccctaatt ccattactac ctgagtgcag agaagatacc agcagtgcgg ttataagaca 5820
tagtgctgat tttggtattg catttgatgg tgattttgat aggtgttttt tctttgatga 5880
aaatggacaa tttattgaag gatactacat tgttggttta ttagcggaag ttttttagg 5940
gaaatatcca aacgcaaaaa tcattcatga tctcgcctt atatggaata ctattgatat 6000

cgtagaaagt catggtggta tacctataat gactaaaacc ggtcatgctt acattaagca 6060
aagaatgcgt gaagaggatg ccgtatatgg cggcgaaatg agtgcgcac attattttta 6120
agattttgca tactgcgata gtggaatgat tccttggatt ttaattttgtg aacttttgag 6180
tctgacaaat aaaaaattag gtgaactggt ttgtggttgt ataaacgact ggccggcaag 6240
tgagaaaata aactgtacac tagacaatcc gcaaaatgaa atagataaat tatttaatcg 6300
ttacaaagat agtgcccttag ctgttgatta cactgatgga ttaactatgg agttctctga 6360
ttggcgtttt aatgtttagat gctcaaatac agaacctgta gtacgattga atgtagaatc 6420
taggaataat gctattctta tgcaggaaaa aacagaagaa attctgaatt ttatatcaaa 6480
ataaatttgc acctgagttc ataatgggaa caagaaatat atgaaagtac ttctgactgg 6540
ctcaactggc atggttggta agaatatatt agagcatgat agtgcaagta aatataatat 6600
acttactcca accagctctg atttgaattt attagataaa aatgaaatag aaaaattcat 6660
gcttatcaac atgccagact gtattataca tgcagcggga ttagttggag gcattcatgc 6720
aaatataagc aggcggtttg attttctgga aaaaaatttg cagatgggtt taaatttagt 6780
ttccgtcgca aaaaaactag gtatcaagaa agtgcttaac ttgggtagtt catgcatgta 6840
cccaaaaaac tttgaagagg ctattcctga gaaagctctg ttaactgggtg agctagaaga 6900
aactaatgag ggatatgcta ttgcgaaaat tgctgtagca aaagcatgcg aatatatata 6960
aagagaaaac tctaattatt ttataaaaac aattatccca tgtaatttat atgggaaata 7020
tgataaattt gatgataact cgtcacatat gattccggca gttataaaaa aaatccatca 7080
tgcgaaaatt aataatgtcc cagagatcga aatttggggg gatggtaatt cgcgcctgta 7140
gtttatgtat gcagaagatt tagctgatct tattttttat gttattccta aaatagaatt 7200
catgccta atggtaaatg ctggtttagg ttacgattat tcaattaatg actattataa 7260
gataattgca gaagaaattg gttatactgg gagtttttct catgatttaa caaaaccaac 7320
aggaatgaaa cggaagctag tagatatctt attgcttaat aaaattgggtt ggtcaagtca 7380
ctttgaactc agagatggca tcagaaaagac ctataattat tacttggaaga atcaaaataa 7440
atgattacat acccacttgc tagtaatact tgggatgaat atgagtatgc agcaatacag 7500
tcagtaattg actcaaaaat gtttaccatg ggtaaaaagg ttgagttata tgagaaaaat 7560
tttgctgatt tgtttggtag caaatatgcc gtaatgggta gctctgggtc tacagcta at 7620
ctgttaatga ttgctgccct tttcttcact aataaaccaa aacttaaaag aggtgatgaa 7680
ataatagtag ctgcagtgtc atggtctacg acatattacc ctctgcaaca gtatggctta 7740
aaggtgaagt ttgtcgatat caataaagaa actttaaaata ttgatatcga tagtttgaaa 7800
aatgctatct cagataaaac aaaagcaata ttgacagtaa atttattagg taatccta at 7860
gattttgcaa aaataaatga gataataaat aatagggata ttatcttact agaagataac 7920
tgtgagtcga tgggcgcggt ctttcaaaat aagcaggcag gcacattcgg agttatgggt 7980
acctttagtt ctttttactc tcatcatata gctacaatgg aagggggctg cgtagttact 8040
gatgatgaag agctgtatca tgtattgttg tgccttcgag ctcatgggtg gacaagaaat 8100
ttacaaaag agaatatggt tacaggcact aagagtgatg atattttcga agagtcgttt 8160
aagtttgttt taccaggata caatgttcgc ccacttgaaa tgagtgggtg tattgggata 8220
gagcaactta aaaagttacc aggttttata tccaccagac gttccaatgc acaatatttt 8280
gtagataaat ttaaagatca tccattcctt gatatacaaa aagaagttgg tgaaagtagc 8340
tggtttgggt tttccttcgt tataaaggag ggagctgcta ttgagaggaa gagtttagta 8400
aataatctga tctcagcagg cattgaatgc cgaccaattg ttactgggaa ttttctcaaa 8460
aatgaacgtg ttttgagtta ttttgattac tctgtacatg atacggtagc aaatgccgaa 8520
tatatagata agaattggtt ttttgtcgga aaccaccaga tacctttggt taatgaaata 8580
gattatctac gaaaagtatt aaaataacta acgaggcact ctatttcgaa tagagtgcct 8640
ttaagatggt attaacagtg aaaaaaattt tagcgtttgg ctattctaaa gtactaccac 8700
cggttattga acagtttgtc aatccaattt gcattctcat tatcacacca ctaatactca 8760
accacctggg taagcaaagc tatggtaatt ggattttatt aattactatt gtatcttttt 8820
ctcagttaat atgtggagga tgttccgcat ggattgcaaa aatcattgca gaacagagaa 8880

ttcttagtga tttatcaaaa aaaaatgctt tacgtcaa at ttcctataat ttttcaattg 8940
ttattatcgc atttgcggtta ttgatttctt ttcttatatt aagtatgtt ttcttcgatg 9000
ttgcgaggaa taattcttca ttcttattcg cgattattat ttgtgggttt tttcaggaag 9060
ttgataattt atttagtggt gcgctaaaag gttttgaaaa atttaagtga tcatgttttt 9120
ttgaagtaat tacaagagtg ctctgggctt ctatagtaat atatggcatt tacggaaatg 9180
cactcttata ttttcatgt ttagccttta ccattaaagg tatgctaaaa tatattcttg 9240
tatgtctgaa tattaccggt tgtttcatca atcctaattt taatagagtt gggattgtta 9300
atttgttaaa tgagtcacaaa tggatgtttc ttcaattaac tgggtggcgtc tcacttagtt 9360
tgtttgatag gctcgtaata ccattgattt tatctgtcag taaactggct tcttatgtcc 9420
cttgccctca actagctcaa ttgatgttca ctctttctgc gtctgcaaat caaatattac 9480
taccaatgtt tgctagaatg aaagcatcta acacatttcc ctctaattgt ttttttaaaa 9540
ttctgcttgt atcactaatt tctgttttgc cttgtcttgc gttattcttt tttggctcgtg 9600
atatattatc aatatggata aaccctacat ttgcaactga aaattataaa ttaatgcaaa 9660
ttttagctat aagttacatt ttattgtcaa tgatgacatc ttttcatttc ttgttattag 9720
gaattggtaa atctaagctt gttgcaaatt taaatctggt tgcagggctc gcacttgctg 9780
cttcaacggt aatcgagct cattatggcc tttatgcaat atctatggta aaaataatat 9840
atccggcttt tcaattttat tacctttatg tagcttttgt ctattttaat agagcgaaaa 9900
atgtctattg atttactttt ttcaattact gaaatcgcaa ttgttttttc ttgcactatt 9960
tacatattta ctcaatgttt gttaatgogg aggatctatt tagataaaaag tattttaatt 10020
cttttatgct tgctcttttt tttagtaatc attcaacttc ctgagcttaa tgtaaacggt 10080
ttggtcgatt ctttaaagtt atcactgctt ttattgatgg tctttatcgc ttttcaaaaa 10140
ccgaaattat gcttggtgggt tattattgca ttgttgtttt tgaactctgc atttaatttt 10200
ttatatttaa agacattcga taagtttagc tcatttcctt ttactttttt tatattgctg 10260
ttttacttgt ttagattggg aattggtaat ttaccggttt ataaaaataa aaaattttac 10320
gcgttgattt ttctctttat attaatagac ataatgcagt cattgttaat aaattatagg 10380
gggcagattt tatattccgt aatttgcatc ctgatacttg tgtttaaagt taatttaaga 10440
aaaaagattc catacttttt tttaatgctg ccagttttat atgtaattat tatggcttat 10500
attggtttta attatttcaa taaaggcgta actttttttg aacctacagc aagtaatatt 10560
gaacgtacgg ggatgatata ttatttggtt tcacagcttg gtgattatat attccatggt 10620
atggggacat taaatttctt aaataacggc ggacaatata agacgttata tggacttcca 10680
tcattaattc ctaatgacce tcatgatttt ttattacggt tctttataag tattggtgtg 10740
ataggagcat tggtttatca ttctatattt ttgttttttt ttaggagaat atctttctta 10800
ttatatgaga gaaatgctcc tttcattggt gtaagttggt tgttactgtt acaagttgtg 10860
ttaatttata cattaaaccc ttttgatgct tttaatcgat tgatttgogg gcttacagtt 10920
ggagttgttt atggatttgc aaaaattaga taagtatacc tgtaatggaa atttagacgc 10980
tccacttggt tcaataatca ttgcaactta taattctgaa cttgatatag ctaagtgttt 11040
gcaatcggtta actaatcaat cttataagaa tattgaaatc ataataatgg atggaggatc 11100
ttctgataaa acgcttgata ttgcaaaatc gtttaaagac gaccgaataa aaatagtttc 11160
agagaaaagat cgtggaattt atgatgcctg gaataaagca gttgatttat ccattgggtga 11220
ttgggtagca ttatttggtt cagatgatgt ttactatcat acagatgcaa ttgcttcatt 11280
gatgaagggg gttatggtat ctaatggcgc cctgtgtggt tatgggagga cagcgcacga 11340
aggtcccgat aggaacatat ctggattttc aggcagtga tgggtacaacc taacaggatt 11400
taagtttaat tattacaaat gtaatttacc attgcccatt atgagcgcaa tatattctcg 11460
tgatttcttc agaaacgaac gttttgatat taaattaaaa attgttgctg acgctgattg 11520
gtttctgaga tgtttcatca aatggagtaa agagaagtca ccttatttta ttaatgacac 11580
gaccttatt gttagaatgg gatatggtgg ggtttcgact gatatttctt ctcaagttaa 11640
aactacgcta gaaagtttca ttgtacgcaa aaagaataat atatcctgtt taaacatata 11700
gctgattctt agatatgcta aaattctggt gatggtagcg atcaaaaaa tttttggcaa 11760

taatgtttat aaattaatgc ataacgggta tcattcccta aagaaaatca agaataaaat 11820
atgaagattg tttatataat aaccgggctt acttgtggtg gagccgaaca ccttatgacg 11880
cagtttagcag accaaatggt tatacgcggg catgatgtta atattatttg tctaactggg 11940
atatctgagg taaagccaac acaaaatatt aatattcatt atgttaatat ggataaaaaat 12000
tttagaagct ttttagagc tttatttcaa gtaaaaaaaa taattgtcgc cttaaagcca 12060
gatataatac atagtcatat gtttcatgct aatattttta gtcgttttat taggatgtgtg 12120
attccagcgg tgcccctgat atgtaccgca cacaacaaaa atgaagggtg caatgcaagg 12180
atgttttgtt atcgactgag tgatttttta gcttctatta ctacaaatgt aagtaaagag 12240
gctgttcaag agtttatagc aagaaaggct acacctaaaa ataaaatagt agagattccg 12300
aattttatta atacaaataa atttgatttt gatattaatg tcagaaagaa aacgcgagat 12360
gcttttaatt tgaaagacag tacagcagta ctgctcgcag taggaagact tgttgaagca 12420
aaagactatc cgaacttatt aaatgcaata aatcatttga ttctttcaaa aacatcaaat 12480
tgtaatgatt ttattttgct tattgtctggc gatggcgcat taagaaataa attattggat 12540
ttggtttgtc aattgaatct tgtggataaa gttttcttct tggggcaaag aagtgatatt 12600
aaagaattaa tgtgtgctgc agatcttttt gttttgagtt ctgagtggga aggttttggg 12660
ctcgttggtg cagaagctat ggcgtgtgaa cgtcccgttg ttgctaccga ttctggtgga 12720
gttaaagaag tcgttggacc tcataatgat gttatccctg tcagtaatca tattctgttg 12780
gcagagaaaa tcgctgagac acttaaaata gatgataacg caagaaaaat aatagggtatg 12840
aaaaatagag aatatattgt ttccaatttt tcaattaaaa cgatagttag tgagtgggag 12900
cgcttatatt ttaaataatt caagcgtaat aatataattg attgaaaata taagtttgta 12960
ctctggatgc aatagtttct ctatgctggt tttttactgg ctccgtattt ttacttatag 13020
ctggattttg ttatatatca gtattaatct gtctcaactt catctagact acattcaagc 13080
cgcgcatgcg tcgcgcggtg actacacctg acaggagtat gtaatgtcca agcaacagat 13140
cggcgctcgc ggtatggcag tgatggggcg caacctggcg ctcaacatcg aaagccgcgg 13200
ttataaccgtc tccatcttca accgctcccg cgagaaaact gaagaagttg ttgccgagaa 13260
cccggataag aaactgggtc cttattacac ggtgaaagag ttctgtagag ctcttgaaac 13320
cccacgtcgt atcctgttaa tggtaaaagc aggggcggga actgatgctg ctatcgattc 13380
cctgaagccg tatctggata aaggcgacat cattattgat ggtggcaaca ccttcttcca 13440
ggacactatc cgtcgtaacc gtgaactgtc cgcggaaggc ttttaacttca tcggtaccgg 13500
cgtgtccggc ggtgaagagg gcgccctgaa aggccatct atcatgccag gtggccagaa 13560
agaagcgtat gagctggttg cgctatcct gaccaagatt gctgcggtg ctgaagatgg 13620
cgaaccatgt ataacttaca tcggtgctga cggtcgggt cactacgtga agatggtgca 13680
caacggtatc gaatatggcg atatgcagct gattgctgaa gcctattctc tgcttaaagg 13740
cggccttaat ctgtctaacc aagagctggc aaccactttt accgagtgga atgaaggcga 13800
gctaagtagc tacctgattg acatcaccaa agacatcttc accaaaaaag atgaagaggg 13860
taaataacctg gttgatgtga tcctggacga agctgcgaac aaaggcaccg gtaaatggac 13920
cagccagagc tctctggatc tgggtgaacc gctgtcgtg atcacgaat ccgtattcgc 13980
tcgctacatc tcttctctga aagaccagcg cattgcggca tctaaagtgc tgtctggtcc 14040
gcaggctaaa ctggctgggtg ataaagcaga gttcgttgag aaagtccgtc gcgcgctgta 14100
cctgggtaaa atcgtctctt atgccaagg cttctctcaa ctgctgcccg cgtctgacga 14160
atacaactgg gatctgaact acggcgaaat cgcgagatc ttccgcgcgg gctgcatcat 14220
tcgtgcgcag ttcctgcaga aaattactga cgctatgct gaaaacaaag gcattgctaa 14280
cctgttgctg gctccgtact tcaaaaatat cgctgatgaa tatcagcaag cgctgcgtga 14340
tgtagtggct tatgctgtgc agaacggtat tccggtaccg accttctctg cagcggtagc 14400
ctactacgac agctaccgtt ctgcggtact gccggcta atctgattcagg cacagcgtga 14460
ttacttcggt gcgcacacgt ataaacgcac tgataaagaa ggtgtgttcc acaccg 14516

<210> 46

<211> 1380

<212> DNA

<213> Escherichia coli

<400> 46

aacaaatctc agtcttctct tagctctgct attgagcgtc tgtcttctgg tctgcgtatt 60
aacagcgcaa aagacgatgc agcaggtcag gcgattgcta accgttttac ggcaaataatt 120
aaagggtctga cccaggcttc ccgtaacgcg aatgatggta tttctgttgc gcagaccact 180
gaagggtgcgc tgaatgaaat taacaacaac ctgcagcgta ttcgtgaact ttctgttcag 240
gcaactaacg gtactaactc tgacagcgat ctttcttcta tccaggctga aattactcaa 300
cgtctggaag aaattgaccg tgtatctgag caaactcagt ttaacggcgt gaaagtcctt 360
gctgaaaata atgaaatgaa aattcagggtt ggtgctaata atggtgaaac catcactatc 420
aatctggcaa aaattgatgc gaaaactctc ggcttggacg gttttaatat cgtggcgcg 480
cagaaagcaa ccggcagtgta cctgatttct aaatttaaag cgacaggtag tgataattat 540
caaattaacg gtactgataa ctatactgtt aatgtagata gtggagtagt acaggataaa 600
gatggcaaac aagtttatgt gagtgtgcg gatggttcac ttacgaccag cagtgatact 660
caattcaaga ttgatgcaac taagcttgca gtggctgcta aagatttagc tcaaggtaat 720
aagattgtct acgaaggtat cgaatttaca aataccggca ctggcgctat acctgccaca 780
ggtaatggtg aattaaccgc caatgttgat ggtaaggctg ttgaattcac tatttcgggg 840
agtgtgata catcaggtag tagtgcaacc gttgcccta cgacagccct atacaaaaat 900
agtgcagggc aattgactgc aacaaaagtt gaaaataaag cagcgacact atctgatctt 960
gatctgaacg ctgccaagaa aacaggaagc acgttagttg ttaacggtgc aacttacgat 1020
gttagtgcag atggtaaaac gataacggag actgcttctg gtaacaataa agtcatgtat 1080
ctgagcaaat cagaagggtg tagcccgatt ctggtaaacy aagatgcagc aaaatcgttg 1140
caatctacca ccaaccgcgt cgaaactatc gacaaagcat tggctaaagt tgacaatctg 1200
cgttctgacc tcgggtgcagt acaaaaaccg ttcgactctg ccatcaccaa ccttggcaac 1260
accgtaacaa acctgtcttc tgcccgtagc cgtatcgaag atgctgacta cgcgaccgaa 1320
gtgtctaaca tgtctcgtgc gcagatcctg caacaagcgg gtacctctgt tctggcacag 1380

<210> 47

<211> 1497

<212> DNA

<213> Escherichia coli

<400> 47

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgtgtctgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt tcacctetaa cattaaaggc 180
ctgactcagg cggcccgtaa cgccaacgac ggtatctccg ttgcgcagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtgtgcgtg aactgacggg acaggccact 300
accggtacta actctgagtc tgatctgtct tctatccagg acgaaattaa atcccgtctg 360
gatgaaattg accgcgtatc tggtcagacc cagttcaacg gcgtgaacgt gctggcaaaa 420
aatggctcca tgaaaatcca ggttggcgca aatgataacc agactatcac tatcgatctg 480
aagcagattg atgctaaaac tcttggcctt gatggtttta gcgttaaaaa taacgataca 540
gttaccacta gtgtccagt aactgctttt ggtgtacca ccacaaacaa tattaactt 600
actggaatta ccctttctac ggaagcagcc actgatactg gcggaactaa ccagcttca 660
attgaggggtg ttatactga taatggtaat gattactatg cgaaaatcac cgggtggatg 720
aacgatggga agtattacgc agtaacagtt gctaatagat gtacagtgc aatggcgact 780

ggagcaacgg caaatgcaac tgtaactgat gcaaatacta ctaaagctac aactatcact 840
 tcaggcggta cacctgttca gattgataat actgcagggt cgcgaactgc caaccttggg 900
 gctgttagct tagtaaaact gcaggattcc aagggtaatg ataccgatac atatgcgctt 960
 aaagatacaa atggcaatct ttacgctgcg gatgtgaatg aaactactgg tgctgtttct 1020
 gttaaaacta ttacctatac tgactcttcc ggtgccgcca gttctccaac cgcgggtcaaa 1080
 ctgggaggag atgatggcaa aacagaagtg gtcgatattg atggtaaaac atacgattct 1140
 gccgatttaa atggcggtaa tctgcaaaca ggtttgactg ctgggtggtga ggctctgact 1200
 gctgttgcaa atggtaaaac cacggatccg ctgaaagcgc tggacgatgc tatcgcatct 1260
 gtagacaaat tccgttcttc cctcgggtgcg gtgcaaaacc gtctggattc cgcgggttacc 1320
 aacctgaaca acaccactac caacctgtct gaagcgcagt cccgtattca ggacgccgac 1380
 tatgcgaccg aagtgtccaa tatgtcgaaa gcgcagatca tccagcaggc cggtaactcc 1440
 gtgttggtcaa aagctaacca ggtaccgcag caggttctgt ctctgctgca ggggttaa 1497

<210> 48

<211> 1695

<212> DNA

<213> Escherichia coli

<400> 48

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
 gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaagggc 180
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
 gcgtgtctg aaatcaacaa caacttacag cgtattcgtg aactgacggt tcaggcttct 300
 accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctg 360
 gacgaaattg accgcgtatc cggtcaaacc cagttcaacg gtgtgaacgt actggcgaaa 420
 gacggttcga tgaaaattca ggttggtgcg aatgacggcc agactatcac tattgatctg 480
 aagaaaattg actctgatac gctggggctg aatggtttta acgttaacgg caaagggtact 540
 attgcgaaca aagcggcaac cattagtgat ctggcggcga cggggggcga tgttactaac 600
 tcaagcaata ttgttgtcac gacaaagttc aatgccttgg atgcagcgac tgcatttagc 660
 aaactcaaag atggtgattc tgttgccgtt gctgtcaga aatatactta taacgcacgc 720
 accaatgatt ttacgacaga aaatacagta gcgacaggca ctgcaacgac agatcttggc 780
 gctactctga aggetgctgc tgggcagagt caatcaggta catatacctt tgcaaatggg 840
 aaagttaact ttgatgttga tgcaagcggg aatatcacta ttggcggcga aaaggcttct 900
 ttggttggtg gagcgctgac tactaacgat cccaccggct cactccagc aacgatgtct 960
 tccctgttta aggcgcgga tgacaaagat gccgctcaat cctcgattga ttttggcggg 1020
 aaaaaatagc aatttgctgg tggcaattct actaatggtg gcggcggtta attcaaagac 1080
 acggtgtctt ctgacgcgct tttggctcag gttaaagcgg atagtactgc taataatgta 1140
 aaaatcacct ttaacaatgg tcctctgtca ttactgcat cgttccaaaa tgggtgtatct 1200
 ggctccgcgg catogaatgc agcctacatt gatagcgaag gcgaactgac aactactgaa 1260
 tcctacaaca caaattattc cgtagacaaa gacacggggg ctgtaagtgt tacagggggg 1320
 agcggtagcg gtaaatacgc cgcaaacgtg ggtgctcagg cttatgtagg tgcagatggg 1380
 aaattaacca cgaatactac tagtaccggc tctgcaacca aagatccact aaatgcgctg 1440
 gatgaggcaa ttgcatccat cgacaaattc cgttcttccc tgggggctat ccagaaccgt 1500
 ctggattccg cagtcaccaa cctgaacaac accactacca acctgtctga agcgcagtc 1560
 cgtattcagg acgccgacta tgcgaccgaa gtgtccaaca tgctgaaagc gcagatcctc 1620
 cagcaggccg gtaactccgt gttggcaaaa gctaaccagg taccgcagca ggttctgtct 1680
 ctgctgcagg gttaa 1695

- 37 -

<210> 49

<211> 1164

<212> DNA

<213> Escherichia coli

<400> 49

```
aacaagaacc agtctgcgct gtcgagttct atcgagcgct tgtcttctgg cttgcgtatt 60
aacagcgcgga aggatgacgc cgcgggtcag gcgattgcta accgttttac ttctaacatt 120
aaaggcctga ctcaggctgc acgtaacgcc aacgacggta tttctgttgc gcagaccacc 180
gaaggcgcgcc tgtccgaaat taacaacaac ttacagcgctg tgcgtgagct gactgttcag 240
gcgaccaccg gtactaactc tgagtctgac ctgtcttcta tccaggacga aatcaaactct 300
cgcttgggaag agattgatcg tgtttcaagt cagactcaat ttaacggcgt gaatgttttg 360
gctaaagatg ggaaaatgaa cattcaggtt ggggcaagtg atggacagac tatcactatt 420
gatctgaaaa agatcgattc atctacacta aacctctcca gttttgatgc taaaaacttg 480
ggcaccagtg ttaaagatgg ggccaccatc aataagcaag tggcagtaga tgctggcgac 540
tttaaagata aagcttcagg atcgttaggt accctaaaat tagttgagaa agacggtaag 600
tactatgtaa atgacactaa aagtagtaag tactacgatg ccgaagtaga tactagtaag 660
ggtgaaaatta acttcaactc taaaaatgaa agtggaacta ctctactgc agcgacggaa 720
gtaactactg ttggcgcgga tgtaaaattg gatgcttctg cacttaaagc caaccaatcg 780
cttgtcgtgt ataaagataa aagcggcaat gatgcttata tcattcagac caaagatgta 840
acaactaatc aatcaacttt caatgccgct aatatcagtg atgctggtgt tttatctatt 900
ggtgcatcta caaccgcgcc aagcaattta acagctgacc cgcttaaggc tcttgatgat 960
gcaattgcat ctgttgataa attccgctct tctctcgggt cggttcagaa ccgtctggat 1020
tctgccattg ccaacctgaa caacaccact accaacctgt ctgaagcgca gtcccgattt 1080
caggacgctg actatgcgac cgaagtgtcc aacatgtcga aagcgcagat tatccagcag 1140
gccggttaact ccgtgctggc aaaa 1164
```

<210> 50

<211> 1818

<212> DNA

<213> Escherichia coli

<400> 50

```
atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt tcacctctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgctaacgat ggtatctctc tggcgcagac cactgaaggc 240
gcactgtctg agattaacaa caacttacia cgtgtgcgtg agttgactgt acaggcgacc 300
accggtacta actctgattc tgacctggct tctattcagg acgaaatcaa atcccgtttg 360
tctgaaattg accgcgtatc cgggcagacc cagttcaacg gcgtgaacgt attgtctaaa 420
gatggctccc tgaaaattca ggttggcgca aatgatggtc agactatctc tatcgacctg 480
aagaaaattg actctgatac tctgggtttg aatggtttca acgttaatgg ttctgggtacc 540
attgcaaaca aagcggccac aatcagtgac ttgactgctc agaaagccgt tgacaacggt 600
aatgggtactt ataaagttac aactagcaac gctgcactta ctgcatctca ggcattaagt 660
aagctgagtg atggcgatac tgtagatatt gcaacctatg ctgggtggtac aagttcaaca 720
gttagttata aatacgacgc agatgcaggt aacttcagtt ataacaatac tgcaaacaaa 780
acaagtgtcg cggctggaac tctggcagat actcttctcc cggcagctgg ccagactaaa 840
```

accggtactt acaaggctgc tactggtgat gttaacttta atgttgacgc aactggtaat 900
ctgacaattg gcggacagca agcctacctg actactgatg gtaaccttac aacaaacaac 960
tccggtgggtg cggctactgc aactcttaaa gagctgttta ctcttgctgg cgatggtaaa 1020
tctctgggga acggcggtac tgctaccgtt actctggata atactacgta taatttcaaa 1080
gctgctgcga acgttactga tgggtgctggt gtcacgctg ctgctggtgt aacttataca 1140
gccactgttt ctaaagatgt cattctggca caactgcaat ctgcaagtca ggcagcagca 1200
accgctaccg acggtgatac tgtcgcaacg atcaactata aatctggtgt catgatcggg 1260
tccgctacct ttaccaatgg taaagggtact gccgatggta tgacttctgg tacaactcca 1320
gtcgtagcta cagggtgctaa agctgtatat gttgatggca acaatgaact gacttccact 1380
gcatcttacg atacgactta ctctgtcaac gcagatacag gcgcagtaaa agtgggtatca 1440
ggtactggta ctggtaaaatt tgaagctggt gctgggtgagg atgcttatgt aagcaaagat 1500
ggcaaatata cgacagaaac caccagtgcg ggcactgcaa ccaaagatcc tttggctgcc 1560
ctggatgctg ctatcagctc catcgacaaa ttccgttcc cctgggtgc tatccagaac 1620
cgtctggatt ccgcagtcac caacctgaac aacaccacta ctaacctgtc tgaagcgcag 1680
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca atatgtcgaa agcgcagatc 1740
atccagcagg ccggtaaact tgtgttggca aaagctaacc aggtaccgca gcaggttctg 1800
tctctgctgc aggggttaa 1818

<210> 51

<211> 1344

<212> DNA

<213> Escherichia coli

<400> 51

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtattttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggt tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgctc 360
gacgaaattg accgcgtttc cggtcagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacggttcga tgaagattca ggttggcgcg aatgacgggc agaccatctc tatcgatttg 480
cagaaaattg attcttcaac gctgggattg aaaggtttct cggtatcagg gaacgcatta 540
aaagtttagc atgcgataac tacagttcct ggtgctaata ctggcgatgc cccggttacg 600
gttaaatttg gtgcgaacga taccgctgct gccgcaatgg ctaaaacatt ggggaataagt 660
gatacatcag gcttgtccct acataacgta caaagcgcgg atggtaaagc gacaggaacc 720
tatgttggtc aatctggtaa tgacttctat tcggcttccg ttaatgctgg tggcgttggt 780
acgcttaata ccaccaatgt tacttttact gatcctgcga acggtgttac cacagcaaca 840
cagacaggtc agcctatcaa ggtcacgacg aatagtgtg gcgcggctgt tggctatggt 900
actattcaag gcaaagatta ccttgctggt gcagacggtg aggatgcaat tgaaaacggt 960
ggtgacgctg caacaaatga agacacaaaa atccaactta ccgatgaact cgatgttgat 1020
ggttctgtaa aaacagcggc aacagcaaca tttcttggtg ctgcaaccaa cgatccgctg 1080
gcacttttag acaaaagctat ctgcgaagtt gatactttcc gtcctccct cgggtgccgt 1140
caaaaccgtc tggattctgc ggtcaccaac ctgaataaca ccaccaacca cctgtctgaa 1200
gcgcagtcgc gtattcagga cgcgcactat gcgaccgaag tgtccaacat gtcgaaagcg 1260
cagatcatcc agcaggcggg taactctgtg ctgtctaaag ctaaccagggt accgcagcag 1320
gttctgtctc tgctgcaggg ttaa 1344

<210> 52 - 39 -

<211> 2599

<212> DNA

<213> Escherichia coli

<400> 52

cttctcttag ctctgctatt gagcgtctgt cttctggtct gcgtattaac agcgcaaaag 60
acgatgcagc aggtcaggcg attgctaacc gttttacggc aaatattaaa ggtctgaccc 120
aggcttcccg taacgcgaat gatggtattt ctggtgcgca gaccactgaa ggtgcgctga 180
atgaaattaa caacaacctg cagcgtatct gtgaactttc tgttcaggca actaacggta 240
ctaactctga cagcgatctt tcttctatcc aggctgaaat tactcaacgt ctggaagaaa 300
ttgaccgtgt atctgagcaa actcagttta acggcgtgaa agtccttgct gaaaataatg 360
aaatgaaaat tcaggttggt gctaattgat gtgaaaccat tgacctgccc ccacgattag 420
atacaacact cagttagtaa cgtcgggaatc ttcattctca gaatgacctt ttctccagcc 480
cgctgcaaat tcagacggtg tctgataatt cagcgtggag tgcggggcggc attcgttata 540
atcctgccgc cagtcattaa taattttcct ggcatgaacg atatcgctga accagtgtct 600
attcaaacat tcatcgcgaa atcgtcgtt aaagctctca ataaatccgt tctgcgttgg 660
cttgcccggc tggattaagc gcaactcaac accatgtctc aaggccatt gatccagtgc 720
acggcaagtg aactccggcc cctggtcagt tcttatcgtc gccggatagc ctcgaaacag 780
tgcaatgctg tccagaatac gcgtgacctg aacgcctgaa atcccaaagg caacagtgc 840
cgtcaggcat tcctttgtga aatcatcgac gcaggtaaga cacttgatcc tgcgaccggt 900
ggaaagtgcg tccatgacga aatccatcga ccaggtcaga ttggggcgcg ccggacggag 960
cagcggcaga cgttctgttg ccagcccttt acgacgtctt ctgcgtttta cgcccaggcc 1020
actgaggtga taaagccggt acacgcgctt atgattaaca tgaagccctt cacggcgcag 1080
caactgccaa atacgacggt agccaaaacg cctgcgctcc agtgccagct cagtgtgcg 1140
cctgataaaa tgcgcatcag cagccggacg gtgagcctca tagcggcagg tgcacaggga 1200
taaacctgta agcctgcagg cagcagcttg cgacagaccg gtcgcatcac acatcaacat 1260
cacggcttcc cgcttctggt ctgtcgtcag tactttcgcc caagagccac ctgaagcgcc 1320
tctttatcca gcatggcttc ggcaagcagc ttcttgagtc tgggtgttctc ttctcaagc 1380
gacttcaggc gcttaacttc aggcacctcc ataccgccat acttcttacg ccaggtgtaa 1440
aacgtggcat cggaaatggc atgcttgccg cagagttcac gggcgggtac ccagcttcg 1500
gcttcgcgga gaatactgat gatctgttcg tcggaaaaac gcttcttcat ggggatgtcc 1560
tcatgtggct tatgaagaca ttactaacat cggggtgtac taatcaacgg ggagcaggtc 1620
accatcacta tcaatctggc aaaaattgat gcgaaaactc tcggcctgga cggttttaat 1680
atcgatggcg cgcagaaagc aaccggcagt gacctgattt ctaaatttaa agcgacaggt 1740
actgataatt atcaaattaa cggtactgat aactatactg ttaatgtaga tagtggagta 1800
gtacaggata aagatggcaa acaagtttat gtgagtgtg cggtatggtt acttacgacc 1860
agcagtgata ctcaattcaa gattgatgca actaagcttg cagtggctgc taaagattta 1920
gctcaaggta ataagattgt ctacgaaggt atcgaattta caataccgg cactggcgct 1980
atacctgcc aaggtaatgg taaattaacc gccaatgttg atggtaaggc tgttgaaattc 2040
actatttcgg ggagtgtga tacatcaggt actagtgcga ccgttgcccc tacgacagcc 2100
ctatacaaaa atagtgcagg gcaattgact gcaacaaaag ttgaaaataa agcagcgaca 2160
ctatctgata ttgatctgaa cgctgccaaag aaaacaggaa gcacgttagt tgttaacggc 2220
gcaacttacg atgttagtgc agatggtaaa acgataacgg agactgcttc tggtaacaat 2280
aaagtcatgt atctgagcaa atcagaaggt ggtagcccga ttctggtaaa cgaagatgca 2340
gcaaaatcgt tgcaatctac caccaaccgg ctcgaaacta tcgacaaagc attggctaaa 2400
gttgacaatc tgcgttctga cctcgggtgca gtacaaaacc gtttcgactc tgccatcacc 2460
aaccttggca acaccgtaaa caacctgtct tctgcccgtg gccgtatcga agatgctgac 2520

tacgcgaccg aagtgtctaa catgtctcgt ggcagatcc tgcaacaagc gggtagctct 2580
gttctggcac aggctaacc 2599

<210> 53

<211> 1245

<212> DNA

<213> Escherichia coli

<400> 53

aacaaaaacc agtctgcgct gtcgacttct atcgagcgcc tctcttctgg tctgcgcatt 60
aacagcgcta aagatgacgc tgcggggccag gcgattgcta accgcttcac ttctaacatc 120
aaaggtctga ctcaggccgc acgtaacgcc aacgacggta tctctctggc gcagaccact 180
gaaggcgcac tgtctgaaat caacaacaac ttgcagcgtg ttctggaact gaccgttcag 240
gccactaccg gtactaactc tgattctgac ctgtcttcaa tccaggacga aatcaaatcc 300
cgtctcgatg aaattgaccg cgtatccggc cagactcagt tcaacggcgt gaacgtactg 360
gcaaaagatg gctcgatgaa aattcagggtc ggtgcaaatg atggtcagac aatcagcatt 420
gatttgcaga agattgatc ttctacttta ggggttaaatg gtttttctgt ttccaaaaat 480
gcagtatctg ttggtgatgc tattactcaa ttgcctggcg agacggcagc cgatgcacca 540
gtaaccatca agtttgatga ttcagtaaaa actgatttaa aactgaccga tgcttcaggg 600
ttaagtctgc ataacctcaa agatgaaaat ggtaatttaa ctaaccagta tgttgtacag 660
aatggcggaa aatcttacgc tgctacagtc gctgccaatg gtaatgttac gctgaacaaa 720
gcaaatgtaa cctacagcga tgcgcgaaac ggtattgata ccgcaacgca gtcaggccag 780
ttagttcagg ttggtgcaga ttctaccggc acgcaaaaag cattcgtgtc tgtccaaggt 840
aaaagctttg gcattgatga cgcgccttg aagaataaca ctggtgatgc taccgctact 900
ccaccgggaa catctgggac aacagttgtc gcagcgtcaa ttcactctgag tacgggcaaa 960
aactctgtag acgtgatgt aacggcttcc actgaattca cagggtgcttc aaccaacgat 1020
ccactgactc tgctggacaa agctatcgca tctgttgata aattccgttc ttctttgggg 1080
gcggtacaga accgtctgag ctccgctgta accaacctga acaacaccac caccaacctg 1140
tctgaagcgc agtcccgtat tcaggacgcc gactatgcga ccgaagtgtc caacatgtcg 1200
aaagcgcaga ttatccagca ggcaggtaac tccgtgctgt ccaaa 1245

<210> 54

<211> 1212

<212> DNA

<213> Escherichia coli

<400> 54

aacaaaaacc agtctgcgct gtcgacttct atcgaacgcc tctcttctgg cctgcgtatt 60
aacagtgcga aagatgacgc tgccgggtcag gcgatagcta accgtttcac ctctaacatt 120
aaaggcctga ctcaggctgc gcgtaacgcc aacgacggta tttctctggc gcagaccaca 180
gaaggtgcgt tgtctgaaat caacaacaac ttgcaacgtg tgctgagatt gaccgttcag 240
gcgacgaccg gtactaactc tgattctgac ctgtcatcta ttcaggacga aatcaaatcc 300
cgtctggatg agattgaccg tgtttccggc cagaccagc tcaacggcgt gaatgtactg 360
gcaaaagacg gttcgatgaa gattcagggt ggccgcaatg atggccagac tattagcatt 420
gatttacaga aaattgactc ttctacatta ggggtgaatg gtttctccgt ttctgtcaa 480
tcacttaacg ttggtgatc aattactcaa attacaggag ccgctgggac aaaacctgtt 540
gggtgtgatt tctactgtgt tgcgaaagat ctgactactg cgacaggtaa aactgtcgtc 600
gtttccagcc tgacgttaca caacacctg gatgcgaaag gggctgccac cgcacagttc 660

gtcgttcaat ccggtagtga tttctactcc gcgtccattg accatgcaag tggatgaagt 720
acgttgaata aagccgatgt cgaatacaaaa gacaccgata atggactaac gactgcagct 780
actcagaaag atcagctgat taaagttgcc gctgactctg acggcgcggc tgcgggatat 840
gtaacattcc agggtaaaaa ctacgctaca acggctccag cggcgcttaa tgatgacact 900
acggcaacag ccacagcgaa caaagttgtt gttgaattat ctacagcaac tccgactgcg 960
cagttctcag gggcttcttc tgctgatcca ctggcacttt tagacaaagc cattgcacag 1020
gttgatactt tccgctcttc cctcgggtgcc gttcaaaacc gtctggactc tgcggtaacc 1080
aacctgaaca acaccaccac caacctgtct gaagcgcagt cccgtattca ggacgccgac 1140
tatgcgaccg aagtgtctaa catgtcgaaa gcgcagatca tccagcaggc gggtaactct 1200
gtgctgtcta aa 1212

<210> 55

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 55

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgtctgtcag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcccgggg tcaggcgatt gctaaccgtt ttacttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaaaattca ggttggtgcg aatgacggtg aaactatcac tatcgacctg 480
aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaaattg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacc 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgcctgattc 660
gataaattag ggaatggcga taaagtcacc gttggcgcg tagattatac ttacaacgct 720
aaatctggtg attttactac caccaaactc actgctggta cgggtgtaga cgccgcgcg 780
caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct taaagccgcg 1020
agcgaaggta gtgacgggtg ctctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggt tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtggtatta ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attgggtactg ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
atccagcagg ccggtaaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggtttaa 1758

<210> 56

<211> 14024

<212> DNA

<213> Escherichia coli

<400> 56

gtaaccaagg gcggtacgtg cataaatttt aatgcttatac aaaactatta gcattaaaaa 60
tatataagaa attctcaaat gaacaaagaa accgtttcaa taattatgcc cgtttacaat 120
ggggccaaaa ctataatctc atcagtagaa tcaattatac atcaatctta tcaagatttt 180
gttttgtata tcattgacga ttgtagcacc gatgatacat tttcattaat caacagtcga 240
tacaaaaaca atcagaaaat aagaatattg cgtaacaaga caaatttagg tgttgcagaa 300
agtcgaaatt atggaataga aatggccacg gggaaatata tttctttttg tgatgcggat 360
gatttgtggc acgagaaaaa attagagcgt caaatcgaag tgttaaataa tgaatgtgta 420
gatgtggtat gttctaatta ttatgttata gataacaata gaaatattgt tggcgaagtt 480
aatgctcctc atgtgataaa ttatagaaaa atgctcatga aaaactacat agggaaattt 540
acaggaatct ataatgcaa caaattgggt aagttttatc aaaaaagat tggtcacgag 600
gattatttga tgtggctgga aataattaat aaaacaaatg gtgctatttg tattcaagat 660
aatctggcgt attacatgcg ttcaaataat tcactatcgg gtaataaaat taaagctgca 720
aaatggacat ggagtatata tagagaacat ttacatttgt cctttccaaa aacattatat 780
tattttttat tatatgcttc aaatggagtc atgaaaaaaa taacacattc actattaagg 840
agaaaggaga ctaaaaagtg aagtcagcgg ctaagttgat tttttattc ctatttacac 900
tttatagtct ccagttgtat ggggttatca tagatgatcg tataacaaat tttgatacaa 960
aggtattaac tagtattata attatatttc agattttttt tgttttatta ttttatctaa 1020
cgattataaa tgaaagaaaa cagcagaaaa aatttatcgt gaactgggag ctaaagttaa 1080
tactcgtttt cctttttgtg actatagaaa ttgctgctgt agttttattt cttaaagaag 1140
gtattcctat atttgatgat gatccagggg gggctaaaact tagaatagct gaaggtaatg 1200
gactttacat tagatatatt aagtattttg gtaatatagt tgtgtttgca ttaattattc 1260
tttatgatga gcataaatc aaacagagga ccatcatatt tgtatatattt acaacgattg 1320
ctttatttgg ttatcgttct gaattgggtg tgctcattct tcaatatata ttgattacca 1380
atatcctgtc aaaggataac cgtaatccta aaataaaaag aataataggg tattttttat 1440
tggtaggggt tgtatgctcg ttgttttata taagtttagg acaagacgga gaacaaaatg 1500
actcatataa taatatgtta aggataatta atagggttaac aatagagcaa gttgaagggtg 1560
ttccatatgt tgtttctgaa tctattaaga acgatttctt tccgacacca gagttagaaa 1620
aggaattaaa agcaataata aatagaatac aggggaataaa gcatcaagac ttattttatg 1680
gagaacgggt acataaacia gtatttgag acatgggagc aaatttttta tcagttacta 1740
cgtatggagc agaactgtta gttttttttg gttttctctg tgtattcatt atccctttag 1800
ggatatatat acctttttat cttttaaaga gaatgaaaaa aacccatagc tcgataaatt 1860
gcgcattcta ttcatatatc attatgattt tattgcaata cttagtggct gggaatgcat 1920
cggccttctt ttttggctct tttctctcgg tattgataat gtgtactoct ctgactttat 1980
tgcatgatac gttaaagaga ttatcacgaa atgaaaatat cagttataac tgtgacttat 2040
aataatgctg aagggttaga aaaaacttta agtagtttat caatttttaa aataaaacct 2100
tttgagatta ttatagttga tggcggctct acagatggaa cgaatcgtgt cattagtaga 2160
tttactagta tgaatattac acatgtttat gaaaaagatg aagggatata tgatgcgatg 2220
aataagggcc gaatgttggc caaaggcgac ttaatacatt atttaaaccg cggcgatagc 2280
gtaattggag atatatataa aaatatcaaa gagccatgtt tgattaaagt tggccttttc 2340
gaaaatgata aacttctggg attttcttct ataaccatt caaatacagg gtattgtcat 2400
caaggggtga ttttccaaa gaatcattca gaatatgatc taaggatata aatatgtgct 2460
gattataagc ttattcaaga ggtgttctct gaagggttaa gatctctatc tttgattact 2520

tccgggttatg taaaatatga tatgggggga gtatcttcaa aaaaaagaat ttttaagagat 2580
aaagagcttg ccaaaattat gtttgaaaaa aataaaaaaa accttattaa gtttattcca 2640
atttcaataa tcaaaatttt attccctgaa cgtttaagaa gagtattgag gaaaatgcaa 2700
tatatttgtc taacttttatt cttcatgaag aatagttcac catatgataa tgaataaaat 2760
caaaaaaata cttaaatttt gcacttttaa aaaatatgat acatcaagtg ctttaggtag 2820
agaacaggaa aggtacagga ttatactctt gtctgttatt tcaagtttga ttagtaaaat 2880
actctcacta ctttctctta tattaactgt aagtttaact ttaccttatt taggacaaga 2940
gagatttggg gtatggatga ctattaccag tcttggtgct gctctgacat ttttgactt 3000
aggtatagga aatgcattaa caaacaggat cgcacattca tttgcgtgtg gcaaaaattt 3060
aaagatgagt cggcaaatta gtggtgggct cactttgctg gctggattat cgtttgtcat 3120
aactgcaata tgctatatta cttctggcat gattgattgg caactagtaa taaaaggat 3180
aaacgagaat gtgtatgcag agttacaaca ctcaattaaa gtctttgtaa tcatatttgg 3240
acttgaatt tattcaaatg gtgtgcaaaa agtttatatg ggaatacaaa aagcctatat 3300
aagtaatat gttaatgcca tatttatatt gttatctatt attactctag taatatcgtc 3360
gaaactacat ggggactac cagttttaat tgcagcact cttggtattc aatacatatc 3420
gggaatctat ttaacaatta atcttattat aaagcgatta ataaagttta caaaagttaa 3480
catacatgct aaaagagaag ctccatattt gatattaaac ggtttttct tttttattt 3540
acagttaggc actctggcaa catggagtgg tgataacttt ataatatcta taacattggg 3600
tgttacttat gtgtgtgtt ttagcattac acagagatta tttcaaatat ctacgggtcc 3660
tcttacgatt tataacatcc cgttatggc tgcttatgca gatgctcatg cagcgaatga 3720
tactcaattt aaaaaaaga cgctcagaac atcattgaaa atagtgggta tttcatcatt 3780
cttattggcc ttcataattag tagtggtcgg tagtgaagtc gttaatattt ggacagaagg 3840
aaagattcag gtacctgaa cattcataat agcttatgct ttatggctctg ttattgatgc 3900
tttttcgaat acatttgcaa gctttttaa tgggttgaa atagttaaac aacaaatgct 3960
tgctgttgta acattgatat tgatcgcaat tccagcaaaa tacatcatag ttagccattt 4020
tgggttaact gttatgttgt actgcttcat ttttatatat attgtaaatt actttatatg 4080
gtataaatgt agttttaa aacatatcga tagacagtta aatataagag gatgaaaatg 4140
aaatatatac cagtttacca accgtcattg acaggaaaag aaaaaagaata tgtaaatgaa 4200
tgtctggact caacgtggat ttcatacaaaa ggaaactata ttcagaagtt tgaaaaataa 4260
tttgcggaac aaaaccatgt gcaatatgca actactgtaa gtaatggaac ggttgctctt 4320
catttagctt tgtagcgtt aggtatatcg gaaggagatg aagttattgt tccaacactg 4380
acatatatag catcagttaa tgctataaaa tacacaggag ccaccccat tttcgttgat 4440
tcagataatg aaacttgga aatgtctgtt agtgacatag acaaaaaat cactaataaa 4500
actaaagcta ttatgtgtgt ccattttatc ggacatccat gtgatatgga acaaatgtga 4560
gaactggcca aaagtagaaa tttgtttgta attgaagatt gcgctgaagc ctttggttct 4620
aaatataaag gtaaatatgt gggaacattt ggagatatat ctacttttag cttttttgga 4680
aataaaacta ttactacagg tgaagggtga atggtgtgca cgaatgacaa aacactttat 4740
gaccgttggt tacattttaa aggccaaagga ttagctgtac ataggcaata ttggcatgac 4800
gttataggct acaattatag gatgacaaat atctgcgctg ctataggatt agcccagtt 4860
gaacaagctg atgattttat atcacgaaa cgtgaaattg ctgatattta taaaaaaat 4920
atcaacagtc ttgtacaagt ccacaaggaa agtaaagatg tttttcacac ttattggatg 4980
gtctcaattc taactaggac cgcagaggaa agagaggaa taaggaatca ccttgcagat 5040
aaactcatg aaacaaggcc agttttttac cctgtccaca cgatgccaat gtactcggaa 5100
aaatatcaaa agcaccctat agctgaggat cttgggtggc gtggaattaa tttacctagt 5160
ttccccagcc tatcgaatga gcaagttatt tatatttgtg aatctattaa cgaattttat 5220
agtataaat agcctaaaat attgtaaagg tcattcatga aaattgcgtt gaattcagat 5280
ggattttacg agtggggcgg tggaattgat tttattaaat atattctgtc aatattagaa 5340
acgaaaccag aaatatgtat cgatattctt ttaccgagaa atgatataca ttctcttata 5400

agagaaaaag catttccttt taaaagtata taaaagcaa ttttaaagag ggaaaggcct 5460
cgatggattt cattaaatag atttaatgag caatactata gagatgcctt tacacaaaat 5520
aatatagaga cgaatcttac ctttattaaa agtaagagct ctgcctttta ttcatttttt 5580
gatagtagcg attgtgatgt tattcttcct tgcattgcgtg ttcccttcggg aaatttgaat 5640
aaaaaagcat ggattgggta tatttatgac tttcaacact gttactatcc ttcatttttt 5700
agtaagcgag aaatagatca aaggaatgtg ttttttaaatt tgatgctcaa ttgcgctaac 5760
aatattattg ttaatgcaca ttcagttatt accgatgcaa ataaatatgt tgggaattat 5820
tctgcaaaac tacattctct tccatttagt ccatgccctc aattaaaatg gttcgtgat 5880
tactctggta atattgccaa atataatatt gacaaggatt attttataat ttgcaatcaa 5940
ttttggaaac ataaagatca tgcaactgct tttagggcatt ttaaaattta tactgaatat 6000
aatcctgatg tttatttagt atgcacggga gctactcaag attatcgatt ccctggatat 6060
tttaatgaat tgatgggttt ggcaaaaag ctcggaattg aatcgaaaat taagatatta 6120
gggcatatac ctaaacttga acaaattgaa ttaatcaaaa attgcattgc tgtaatacaa 6180
ccaaccttat ttgaaggcgg gcctggaggg ggggtaacat ttgacgctat tgcattaggg 6240
aaaaaagtta tactatctga catagatgtc aataaagaag ttaattgcgg tgatgtatat 6300
ttcttttcagg caaaaaacca ttattcatta aatgacgcga tggtaaaagc tgatgaatct 6360
aaaatttttt atgaacctac aactctgata gaattgggtc tcaaaagacg caatgcgtgt 6420
gcagattttc ttttagatgt tgtgaaacaa gaaattgaat cccgatctta atatattcaa 6480
gaggatatata atgactaaag tcgctcttat tacagggtga actggacaag atggatctta 6540
tctagctgag tttttgcttg ataaagggtg tgaagtccat ggtatcaaac gccgagcctc 6600
atcttttaat acagaacgca tagaccatat ttatcaagat ccacatgggt ctaacccaaa 6660
ttttcacttg cactatggag atctgactga ttcactaac ctactagaa ttctaaggga 6720
ggtacagcca gatgaagtat ataatttagc tgctatgagt cacgtagcag tttcttttga 6780
gtctccagaa tatacagccg atgtcgatgc aattgggtaca ttacgtttac tggaagcaat 6840
tcgcttttta ggattggaaa acaaaacgcg tttctatcaa gcttcaacct cagaattata 6900
tggacttggt caggaaatcc ctcaaaaaga atccaccctt ttttatcctc gttcccctta 6960
tgcagttgca aaactttacg catattggat cacggtaaat tatcgagagt catatgggat 7020
ttatgcatgt aatggatatat tgttcaatca tgaatctcca cgccgtggag aaacgtttgt 7080
aacaaggaaa attactcgag gacttgcaaa tattgcacaa ggcttggaat catgtttgta 7140
tttaggggaat atggattcgt tacgagattg gggacatgca aaagattatg ttagaatgca 7200
atgggtgatg ttacaacagg agcaaccgga agattttgtg attgcaacag gagtccaata 7260
ctcagtcctg cagtttgcg aaatggcagc agcacaactt ggtattaaga tgagctttgt 7320
tggtaaaggga atcgaagaaa aaggcattgt agattcgggt gaaggacagg atgctccagg 7380
tgtgaaacca ggtgatgtca ttgttgcgtg tgatcctcgt tatttccgac cagctgaagt 7440
tgatactttg cttggagatc cgagcaaagc taatctcaaa cttggttgga gaccagaaat 7500
tactcttgct gaaatgattt ctgaaatggt tgccaaagat cttgaagccg ctaaaaaaca 7560
ttctctttta aaatcgcatg gtttttctgt aagcttagct ctggaatgat gatgaataag 7620
caacgtattt ttattgctgg tcaccaagga atgggtggat cagctattac ccgacgcctc 7680
aaacaacgtg atgatgttga gttgggttta cgtactcggg atgaattgaa cttgttggat 7740
agtagcgctg ttttggattt tttttcttca cagaaaatcg accaggttta tttggcagca 7800
gcaaaagtcg gaggtatttt agctaacagt tcttatcctg ccgattttat atatgagaat 7860
ataatgatag aggcgaatgt cattcatgct gccacaaaa ataatgtaaa taaactgctt 7920
ttctcgggtt cgtcgtgtat ttatcctaag ttagcacacc aaccgattat ggaagacgaa 7980
ttattacaag ggaaacttga gccacaaat gaaccttatg ctatcgcaaa aattgcaggt 8040
attaaattat gtgaatctta taaccgtcag tttgggcgtg attaccgttc agtaatgcca 8100
accaatcttt atggtccaaa tgacaatttt catccaagta attctcatgt gattccggcg 8160
cttttgcgcc gctttcatga tgctgtggaa aacaattctc cgaatgttgc tgtttgggga 8220
agtgggtact caaagcgtga attcttacat gtagatgata tggcttctgc aagcatttat 8280

gtcatggaga tgccatacga tatatggcaa aaaaatacta aagtaatggt gtctcatatc 8340
aatattggaa caggtattga ctgcacgatt tgtgagcttg cggaacaat agcaaaagtt 8400
gtaggttata aagggcatat tacgttcgat acaacaaagc ccgatggagc ccttcgaaaa 8460
ctacttgatg taacgcttct tcatcaacta ggttggaatc ataaaattac ccttcacaag 8520
ggtcttgaaa atacatacaa ctggtttctt gaaaaccaac ttcaatatcg ggggtaataa 8580
tgtttttaca ttcccaagac tttgccacaa ttgtaaggte tactcctctt atttctatag 8640
atgtgattgt ggaaaacgag tttggcgaaa ttttgctagg aaaacgaatc aaccgcccgg 8700
cacagggcta ttggttcggt cctggtggta ggggtgttgaa agatgaaaaa ttgcagacag 8760
cctttgaacg attgacagaa attgaactag gaattcggtt gcctctctct gtgggtaagt 8820
tttatgggat ctggcagcac ttctacgaag acaatagtat ggggggagac ttttcaacgc 8880
attatatagt tatagcattc cttcttaaat tacaaccaa cttttgaaa ttaccgaagt 8940
cacaacataa tgcttattgc tggctatcgc gagcaaagct gataaatgat gacgatgtgc 9000
attataattg tcgcgcatat ttaacaata aaacaaatga tgcgattggc ttagataata 9060
aggatataat atgtctgatg cgccaataat tgctgtagtt atggcgggtg gtacaggcag 9120
tcgtctttgg ccactttctc gtgaactata tccaaagcag tttttacaac tctctggtga 9180
taacaccttg ttacaaacga ctttgctacg actttcaggc ctatcatgtc aaaaaccatt 9240
agtataaca aatgaacagc atcgctttgt tgtggctgaa cagttaaggg aaataaataa 9300
attaaatggt aatattattc tagaaccatg cgggcgaaat actgcaccag caatagcgat 9360
atctgcgttt catgcgttaa aacgtaatcc tcaggaagat ccattgcttc tagttcttgc 9420
ggcagaccac gttatagcta aagaaagtgt tttctgtgat gctattaaaa atgcaactcc 9480
catcgcta at caaggtaaaa ttgtaacggt tgggaattata ccagaatatg ctgaaactgg 9540
ttatgggtat attgagagag gtgaactatc tgtaccgctt caagggcatg aaaatactgg 9600
tttttattat gtaataaagt ttgtcgaaaa gcctaatacgt gaaaccgcag aattgtatat 9660
gacttctggt aatcactatt ggaatagtgg aatattcatg ttttaaggcat ctgtttatct 9720
tgaggaattg agaaaattta gacctgacat ttacaatggt tgtgaacagg ttgcctcatc 9780
ctcatacatt gatctagatt ttattcgatt atcaaaagaa caatttcaag attgtcctgc 9840
tgaatctatt gatttttgctg taatggaaaa aacagaaaaa tgtgttgat gcctgttgga 9900
tattggttgg agtgacgttg gatcttgcca atcgttatgg gacattagtc taaaatcgaa 9960
aacaggagat gtatgtaaag gtgatataat aacctatgat actaagaata attatatcta 10020
ctctgagtca gcgttggtag ccgccattgg aattgaagat atggttatcg tgcaaaactaa 10080
agatgccgtt cttgtgtcta aaaagagtga tgtacagcat gtaaaaaaaa tagtcgaaat 10140
gcttaaattg cagcaacgta cagagtatat tagtcatcgt gaagttttcc gaccatgggg 10200
aaaatttgat tcgattgacc aaggtgagcg atacaaagtc aagaaaatta ttgtgaaacc 10260
tggtgagggg ctttctttaa ggatgcatca ccacgttct gaacattgga tcgtgcttct 10320
tggtacagca aaagtaaccc ttggcgataa aactaaacta gtcaccgcaa atgaatcgat 10380
atacattccc cttggcgag cgtatagtct tgagaatccg ggcataatcc ctcttaatct 10440
tattgaagtc agttcagggg attatttggg agaggatgat attataagac agaaagaacg 10500
ttacaaacat gaagattaac atatgaaatc tttaacctgc tttaaagcct atgatattcg 10560
cgggaaatta ggcgaagaac tgaatgaaga tattgcctgg cgcattgggc gtgcctatgg 10620
cgaatttctc aaaccgaaaa ccattgtttt aggcggtgat gtccgcctca ccagcgaagc 10680
gttaaaactg gcgcttgca aaggtttaca ggatgcgggc gtcgatgtgc tggatatcgg 10740
tatgtccggc accgaagaga tctatttcgc cacgttccat ctcgagtggt atggcggcat 10800
cgaagttacc gccagccata acccgatgga ttacaacggc atgaagctgg tgcgcgaagg 10860
ggctcgccc atcagcggtg ataccggact gcgcgatgtc cagcgtctgg cagaagccaa 10920
tgacttccct cctgtcgatg aaaccaaagc tggctgctat cagcaaatca atctgcgtga 10980
cgcttacggt gatcacctgt tcggttatat caacgtcaa aacctcacgc cgctcaagct 11040
gggtgatcaac tccgggaacg ggcagcggg tccggtggtg gacgccattg aagcccgatt 11100
taaagccctc ggcgcaccgg tgggaattaat caaagtacac aacacgccg acggcaattt 11160

ccccaacggg attcctaacc cgctgctgcc ggaatgccgc gacgacaccc gtaatgcggt 11220
catcaaacac ggcgcgata tgggcattgc ctttgatggc gattttgacc gctgtttcct 11280
gtttgacgaa aaagggcagt ttatcgaggg ctactacatt gtcggcctgc tggcagaagc 11340
gttcctcgaa aaaaatcccg gcgcgaagat catccacgat ccacgtctct cctggaacac 11400
cgttgatgtg gtgactgccg caggcggcac cccggtaatg tcgaaaaccg gacacgcctt 11460
tattaaagaa cgtatgcgca aggaagacgc catctacggt ggcgaaatga gcgctcacca 11520
ttacttccgt gatttcgctt actgcgacag cggcatgac cctgggctgc tggtcgccga 11580
actggtgtgc ctgaaaggaa aaacgctggg cgaaatggtg cgcgaccgga tggcggcggt 11640
tccggcaagc ggtgagatca acagcaaact ggcgcaaccc gttgaggcaa ttaatcgct 11700
ggaacagcat tttagccgcg aggcgctggc ggtggatcgc accgatggca tcagcatgac 11760
ctttgccgac tggcgcttta acctgcgctc ctccaacacc gaaccgggtg tgcggttgaa 11820
tgtggaatca cgcggtgatg taaagctaact ggaaaagaaa actaaagctc ttcttaaatt 11880
gctaagtgag tgattattta cattaatcat taagcgtatt taagattata ttaaagtaat 11940
gttattgcgg tatatgatga atatgtgggc ttttttatgt ataacgacta taccgcaact 12000
ttatctagga aaagattaat agaaataaag ttttgactg accaatttgc atttcacgtc 12060
acgattgaga cgttcctttg cttaagacat tttttcatcg cttatgtaat aacaaatgtg 12120
ccttatataa aaaggagaac aaaatggaac ttaaaataat tgagacaata gatttttatt 12180
atccctgttt acgatattat agccaaagtt gtatcctgca tcagtcctgc aatatttcac 12240
gagtgccttg ttaactgaat acatgtctgc cattttccag atgataacga cgtcatcgca 12300
attgatggta aaacacttcg gcacacttat gacaagagtc gtcgcagagg agtggttcat 12360
gtcattagtgc cgtttcagca atgcacagtc tggctcctcg atagatcaag acggatgaga 12420
aacctaatgc gttcacagtt attcatgaac tttctaaaat gatgggtatt aaaggaaaaa 12480
taatcataac tgatgcgatg gcttgccaga aagatattgc agagaagata taaaaacaga 12540
gatgtgatta tttattcgct gtaaaaggaa ataagagtcg gcttaataga gtctttgagg 12600
agatatttac gctgaaagaa ttaaaataat caaaacatga cagttacgca attagtga 12660
agaggcacgg cagagacgat gtccgtcttc atattgtttg agatgctcct gatgagctta 12720
ttgatttcac gtttgaatgg aaagggctgc agaatttatg aatggcagtc cactttctct 12780
caataatagc agagcaaaag aaagaatccg aatgacgat caaatattat attagatctg 12840
ctgctttaac cgcagagaag ttcgccacag taaatcgaaa tcaactggcg atggagaata 12900
agttgcacag tagcctgatg tggtaatgaa tgaaatcgac tataatataa gaaggcgagt 12960
tgcattcgaa tgattttcta gaatgcggca catcgctatt aatatctgac aatgataatg 13020
tattcaaggc aggattatca tgtaagatgc gaaaagcagt catggacaga aacttcctag 13080
cgtcaggcat tgcagcgtgc gggctttcat aatcttgc atggttttgat aagatatttc 13140
tttgagatg ggaaaatgaa tttgatggt atttttggtg ctggaagtta tggtagagaa 13200
acaataccca ttctaaatca acaataaaag caagaatgtg gttctgacta tgctctggtt 13260
tttggtgatg atgttttggc aggaagaaa gttaatggtt ttgaagtgc ttcaaccaac 13320
tgctttctaa aagccctta tttaaaaaag tattttaatg ttgctattgc taatgataag 13380
atacgacaga gagtgtctga gtcaatatta ttacacgggg ttgaaccaat aactataaaa 13440
catccaaata gcgtgttcta tgatcatact atgataggta gtggcgctat tatttctccc 13500
tttgttacaa tatctactaa tactcatata gggaggtttt ttcatgcaa catatactca 13560
tacgttgac atgattgtca aataggagac tatgttacat ttgctcctgg ggctaaatgt 13620
aatggatatg ttgttattga agacaatgca tatataggct cgggtgcagt aattaagcag 13680
ggtgttctta atcgccact tattattggc gcgggagcca ttataggat ggggctgtt 13740
gtcactaaaa gtgttctgc cgggtataact gtgtgcgaa atccagcaag agaaatgaaa 13800
agatcgccaa catctattta atgggaatgc gaaaacacgt tccaaatggg actaatgttt 13860
aaaatatata taatttcgct aatttactaa attatggctt ctttttaagc tatcctttac 13920
ttagttatta ctgatacagc atgaaattta taatactctg atacattttt atacgttatt 13980
caagccgcat atctagcggg aaccctgac aggagtaa aatg 14024

- 47 -

<210> 57

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 57

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaagagc 180
ctgactcagg cggcccgtaa cgccaacgac ggtatttctg ttgcgagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggg tcaggccact 300
acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360
gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacggttcaa tgaaaattca ggttggtgag aatgacggcg aaaccatcac gatcgacctg 480
aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaattg taaaggtact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacc 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcaca gttggcggcg tagattatac ttacaacgct 720
aaatctgggtg attttactac cactaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctgctg attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtctgt caaagcagcg 1020
agcgaaggta gtgacgggtc ctctctgaca ttcaatggca cagaatatac catcgcaaaa 1080
gcaactcctg cgacaaccac tccagtagct ccgttaatcc ctgggtgggag tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtggatca ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtagct ctgtaaatgt gaactccgcg 1440
ggtaaatca ctactgagac taccagtgtt ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg caatcagctc catcgacaaa ttcggttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc cgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1680
attcagcagg ccggtaaactc cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggtttaa 1758

<210> 58

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 58

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcagcggg tcaggcgatt gctaaccgtt ttacttctaa cattaagagc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcgagac caccgaaggc 240

gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggg tcaggccact 300
acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360
gatgaaattg accgcgatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacggttcaa tgaaaattca ggttggtgcg aatgacggcg aaaccatcac gatcgacctg 480
aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaatgg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaaccacc 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcaca gttggcgggc tagattatac ttacaacgct 720
aaatctggtg attttactac cactaaatct actgctggta cgggtgtaaa cgccgcggcg 780
caggctgctg attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct caaagcagcg 1020
agcgaaggta gtgacggtag ctctctgaca ttcaatggca cagaatatac catcgcaaaa 1080
gcaactcctg cgacaaccac tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggtgcccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgctgg tgaatccagt 1260
gatgctgoga agtcttatgt ggatgataaa ggtggtatca ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggcactg ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg caatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcgggtcac caacctgaac aacaccacta ccaacctgtc cgaagcgag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcgagtc 1680
atccagcagg ccggtaactc cgtgctggca aaagctaacc aggtaccgca gcagggttctg 1740
tctctgctgc aggggttaa 1758

<210> 59

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 59

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgcgcgggg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggg tcagggttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
gacgaaattg accgcgatc cggctcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaaaattca ggttggtgcg aatgacggg aaactatcac tatcgacctg 480
aagaaaaatcg attctgatac tctgggtctg aatgggttta acgtaaatgg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaaccacc 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcacc gttggcgggc tagattatac ttacaacgct 720
aaatctggtg attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900

```

tcagcaggta atatcaccat cgggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgctgct taaagccgcg 1020
agcgaaggta gtgacgggtgc ctctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtggtatta ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtactg ctgtaaagt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcg 1620
tcccgtattc aggacgcgca ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
atccagcagg ccggttaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggtttaa

```

1758

<210> 60

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 60

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg cggcccgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggg tcaggccact 300
acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360
gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacggttcaa tgaaaattca ggttggtgcg aatgacggcg aaaccatcac gatcgacctg 480
aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaatgg taaaggact 540
attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaacc 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcaca gttggcggcg tagattatac ttacaacgct 720
aaatctgggtg attttactac cactaaatct actgctggtg cgggtgtaga cgccgcggcg 780
caggctgctg attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cgggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgctgct caaagcagcg 1020
agcgaaggta gtgacgggtgc ctctctgaca ttcaatggca cagaatatac catcgcaaaa 1080
gcaactcctg cgacaaccac tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtggtatca ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtactg ctgtaaagt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg caatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560

```

cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc cgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1680
attcagcagg ccggtaaact cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa 1758

<210> 61

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 61

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatctctg ttgctcgagc caccgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtattcgtg aactgacggt tcaggccact 300
acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360
gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
gacggttcaa tgaaaattca ggttggtgag aatgacggcg aaaccatcac gatcgacctg 480
aaaaaaatcg attctgatac tctgggtctg aatggcttta acgtaaattg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgtt ttaacttctg ctggcgcgaa gttaaaccac 600
acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcaca gttggcgagg tagattatac ttacaacgct 720
aaatctgggt attttactac cactaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctgctg attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatgggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cgggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgctc caaagcagcg 1020
agcgaaggta gtgacgggtc ctctctgaca ttcaatggca cagaatatac catcgcaaaa 1080
gcaactcctg cgacaaccac tccagtagct ccgttaatcc ctgggtggat tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtggtatca ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggcactg ctgtaaattg gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgetgcc 1500
ctggacgacg caatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcggtcac caacctgaac aacaccacta ccaacctgtc cgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1680
atccagcagg ccggtaaact cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa 1758

<210> 62

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 62

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
 gcgaaggatg acgccgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
 ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
 gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggcttct 300
 accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
 gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
 gacggttcga tgaaaattca ggttggtgag aatgacggtg aaactatcac tatcgacctg 480
 aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaaattg taaagggtact 540
 attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaaccacc 600
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
 gataaattag ggaatggcga taaagtcacc gttggcggcg tagattatac ttacaacgct 720
 aaatctgggtg attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780
 caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
 ggtaaattctg ttaatggttc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
 tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
 acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtctgt taaagccgcg 1020
 agcgaaggta gtgacggtgc ctctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
 gcaactctctg cgacaacctc tccagtagct ccgttaatcc ctgggtggat ttcttatcag 1140
 gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
 attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
 gatgctgcga agtcttatgt ggatgataaa ggtgggtatta ctaacgttgc cgactataca 1320
 gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
 actgatacca ataaagatta tgctccagca attgggtactg ctgtaaattg gaactccgcg 1440
 ggtaaaatca ctactgagac taccagtgtc ggttctgcaa cgaccaaccc gcttgctgcc 1500
 ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
 cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620
 tcccgattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
 atccagcagg ccggtaaactc cgtgctggca aaagccaacc aggtaccgca gcagggtctg 1740
 tctctgctgc aggggttaa 1758

<210> 63

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 63

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
 aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
 gcgaaggatg acgccgcagg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
 ctgactcagg cggcccgtaa cgccaacgac ggtatttctg ttgcgcagac caccgaaggc 240
 gcgctgtccg aaatcaacaa caacttacag cgtattctgt aactgacggt tcaggccact 300
 acagggacta actccgattc tgacctggac tccatccagg acgaaatcaa atctcgtctt 360
 gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt gctggcgaaa 420
 gacggttcga tgaaaattca ggttggtgag aatgacggcg aaaccatcac gatcgacctg 480
 aaaaaaatcg attctgatac tctgggtctg aatgggttta acgtaaattg taaagggtact 540
 attaccaaca aagctgcaac ggtaagtgat ttaacttctg ctggcgcgaa gttaaaccacc 600
 acgacaggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660

```

gataaattag ggaatggcga taaagtcaca gttggcggcg tagattatac ttacaacgct 720
aaatctggtg attttactac cactaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctgctg attcagcttc aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaactctg ttaatgggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct caaagcagcg 1020
agcgaaggta gtgacgggtg ctctctgaca ttcaatggca cagaatatac catcgcaaaa 1080
gcaactcctg cgacaaccac tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtggtatca ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtactg ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtt ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg caatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc cgaagcgcag 1620
tcccgatttc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatc 1680
attcagcagg ccggttaactc cgtgctggca aaagctaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa

```

1758

<210> 64

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 64

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcggg tcaggcgatt gctaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac caccgaaggc 240
gcgctgtctg aaatcaacaa caacttacag cgtatccgtg agctgacggg tcaggcttct 300
accggaacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgctt 360
gatgaaattg accgcgtatc cggccagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaataattc ggttggtgcg aatgacgggt aaactatcac tatcgacctg 480
aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaaatgg taaagggtact 540
attaccaaca aagctgcaac ggtaagtgtt ttaacttctg ctggcgcgaa gttaaacc 600
acgacagggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcacc gttggcggcg tagattatac ttacaacgct 720
aaatctggtg attttactac caccaaatct actgctggta cgggtgtaga cgccgcggcg 780
caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaactctg ttaatgggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct taaagccgcg 1020
agcgaaggta gtgacgggtg ttctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctggtgggat tacttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcga agtcttatgt ggatgataaa ggtggtatta ctaacgttgc cgactataca 1320

```

gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtagct ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtt ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
atccagcagg ccggttaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa 1758

<210> 65

<211> 1758

<212> DNA

<213> Escherichia coli

<400> 65

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgtgtgtcag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcggg tcaggcgatt gtaaccgtt ttacttctaa cattaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatttctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggt tcaggcttct 300
accgggacta actctgattc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacggttcga tgaaaattca ggttggtgag aatgacgggt aaactatcac tatcgacctg 480
aagaaaatcg attctgatac tctgggtctg aatgggttta acgtaaatgg taaaggtagt 540
attaccaaca aagctgcaac ggtaagtgtt ttaacttctg ctggcgcgaa gttaaaccac 600
acgacagggtc tttatgatct gaaaaccgaa aataccttgt taactaccga tgctgcattc 660
gataaattag ggaatggcga taaagtcacc gttggcggcg tagattatac ttacaacgct 720
aaatctggtg attttactac caccaaactc actgctggta cgggtgtaga cgccgcggcg 780
caggctactg attcagctaa aaaacgtgat gcgttagctg ccacccttca tgctgatgtg 840
ggtaaatctg ttaatgggtc ttacaccaca aaagatggta ctgtttcttt cgaaacggat 900
tcagcaggta atatcaccat cgggtggaagc caggcatacg tagacgatgc aggcaacttg 960
acgactaaca acgctggtag cgcagctaaa gctgatatga aagcgtgct taaagccgcg 1020
agcgaaggta gtgacgggtc ctctctgaca ttcaatggca ctgaatatac tatcgcaaaa 1080
gcaactcctg cgacaacctc tccagtagct ccgttaatcc ctggtgggat ttcttatcag 1140
gctacagtga gtaaagatgt agtattgagc gaaaccaaag cggctgccgc gacatcttca 1200
attaccttta attccggtgt actgagcaaa actattgggt ttaccgcggg tgaatccagt 1260
gatgctgcca agtcttatgt ggatgataaa ggtggtatta ctaacgttgc cgactataca 1320
gtctcttaca gcgttaacaa ggataacggc tctgtgactg ttgccgggta tgcttcagcg 1380
actgatacca ataaagatta tgctccagca attggtagct ctgtaaatgt gaactccgcg 1440
ggtaaaatca ctactgagac taccagtgtt ggttctgcaa cgaccaaccc gcttgctgcc 1500
ctggacgacg ctatcagctc catcgacaaa ttccgttctt ccctgggtgc tatccagaac 1560
cgtctggatt ccgcagtcac caacctgaac aacaccacta ccaacctgtc tgaagcgcag 1620
tcccgtattc aggacgccga ctatgcgacc gaagtgtcca acatgtcgaa agcgcagatt 1680
atccagcagg ccggttaactc cgtgctggca aaagccaacc aggtaccgca gcaggttctg 1740
tctctgctgc aggggttaa 1758

<210> 66

<211> 1788

- 54 -

<212> DNA

<213> Escherichia coli

<400> 66

```

atggcacaag tcattaatac caacagcctc tcgctgatca ctcaaaataa tatcaacaag 60
aaccagtctg cgctgtcgag ttctatcgag cgtctgtctt ctggcttgcg tattaacagc 120
gcgaaggatg acgccgcggg tcaggcgatt gctaaccgtt ttactttctaa cattaaaggc 180
ctgactcagg ctgcacgtaa cgccaacgac ggtatctctg ttgcacagac cactgaaggc 240
gcgctgtccg aaatcaacaa caacttacag cgtatccgtg agctgacggg tcaggcttct 300
accgggacta actctgatcc ggatctggac tccattcagg acgaaatcaa atcccgtctc 360
gacgaaattg accgcgtatc cggtcagacc cagttcaacg gcgtgaacgt actggcaaaa 420
gacgggttcga tgaaaattca ggtagggtgcg aacgacggcc agactatcac tattgatctg 480
aagaaaattg actctgatac gctggggctg aatgggttta acgtgaatgg ttccggtagc 540
atagccaata aagcggcgac cattagcgac ctgacagcag cgaaaatgga tgctgcaact 600
aatactataa ctacaacaaa taatgcgctg actgcatcaa aggcccttga tcaactgaaa 660
gatggtgaca ctgttactat caaagcagat gcagctcaaa ctgccacggg ctatacatac 720
aatgcatctg ctggtaactt ctcattcagt aatgtatcga ataatacttc agcaaaagca 780
ggtgatgtag cagctagcct tctcccgcg gctgggcaaa ctgctagtgg tgtttacaaa 840
gcagcaagcg gtgaagtga ctttgatggt gatgcgaatg gtaaaattac aatcggagga 900
caggaagcct atttaactag tgatggtaac ttaactacaa acgatgctgg tgggtgcgact 960
gcggctacgc ttgatggttt attcaagaaa gctgggtgatg gtcaatcaat cgggtttaat 1020
aagactgcat cagtcacgat ggggggaaca acttataact ttaaaacggg tgctgatgct 1080
ggtgctgcaa ctgctaacgc aggggtatcg ttcactgata cagctagcaa agaaaccgtt 1140
ttaaataaag tggctacagc taaacaaggc acagcagttg cagctaacgg tgatacatcc 1200
gcaacaatta cctataaatc tggcgttcag acgtatcagg cggtatctgc cgcagggtgac 1260
ggtactgcta gcgcaaaata tgccgataat actgacgttt ctaatgcaac agcaacatac 1320
acagatgctg atgggtgaaat gactacaatt ggttcataca ccacgaagta ttcaatcgat 1380
gctaacaacg gcaaggtaac tgttgattct ggaactggtt cgggtaaaata tgcgccgaaa 1440
gtcggggctg aagtatatgt tagtgctaat ggtactttta caacagatgc aactagcgaa 1500
ggcacagtaa caaagatcc actgaaagct ctggatgaag ctatcagctc catcgacaaa 1560
ttccgttcat ccctgggggc tatccaaaac cgtttggatt ccgccgtcac caacctgaac 1620
aacaccacta ccaacctgtc tgaagcgcag tcccgtattc aggacgccga ctatgcgacc 1680
gaagtgtcca acatgtcgaa agcgcagatt atccagcagg ccggtaaactc cgtgctggca 1740
aaagccaacc aggtaccgca gcaggttctg tctctactgc agggttaa 1788

```

<210> 67

<211> 1398

<212> DNA

<213> Escherichia coli

<400> 67

```

aacaaatctc agtcttctct tagctctgct attgagcgtc tgtcttctgg tctgcgtatt 60
aacagcgcaa aagacgatgc agcaggtcag gcgattgcta accgttttac ggcaaatatt 120
aaaggctctga cccaggcttc ccgtaacgca aatgatggta tttctgttgc gcagaccact 180
gaagggtgcgc tgaatgaaat taacaacaac ctgcagcgta ttctgtaact ttctgttcag 240
gcaactaacg gtactaactc tgacagtgc ctgacctcca tccagtccga aatccagcag 300
cgtctgagtg aaattgaccg tgtttctggg cagactcagt ttaacggcgt taaagtgtct 360
gcttctgatc aggatatgac tattcagggtt ggtgcaaacg acggcgaaac aattactatt 420

```

aaactgcagg aaattaattc cgacacactg ggattatctg gttttggtat taaagatcct 480
 actaaattaa aagccgcaac ggctgaaaca acctattttg gatcgacagt taagcttgct 540
 gacgctaata cacttgatgc agatattaca gctacagtta aaggcactac gactccgggc 600
 caacgtgacg gtaatattat gtctgatgct aacggtaagt tgtacgttaa agttgccggg 660
 tcagataaac ccgctgaaaa tggttattat gaagttactg tggaggatga tccgacatct 720
 cctgatgcag gtaagctgaa gctgggggct ctacgggta ccagcctca agctggtaat 780
 ttaaaggaag tcacaacggg gaaaggggag ggggctattg atgttcagtt gggactgat 840
 accgcaaccg cttctatcac aggtgcaaaa ctctttaagt tagaagacgc caatggcaaa 900
 gatactgggt catttgcggt gattgggtgat gacggtaaac agtatgcagc gaatgttgat 960
 cagaaaacag gagcagtttc cgttaaaaca atgtcttaca ctgatgctga cgggtgcaaa 1020
 cacgacaatg ttaaagttga actgggtgga agcgatggca aaaccgaagt tgtaactgca 1080
 accgatggca aaacttacag tgttagtgat ttacaaggta agagcctgaa aactgattct 1140
 attgcagcaa tttctacgca gaaaacagaa gatccttttg ctgctatcga taaagcactg 1200
 tctcaggttg actcggtgcg ttctaacctt ggtgcaattc aaaatcggtt cgactctgcc 1260
 atcaccaacc ttggcaaacac cgtaaacac ctgtcttctg cccgtagccg tatcgaagat 1320
 gctgactacg cgaccgaagt gtctaacatg tctcgtgccc agatcctgca acaagcgggt 1380
 acctctgttc tggcgagcag 1398

<210> 68

<211> 1479

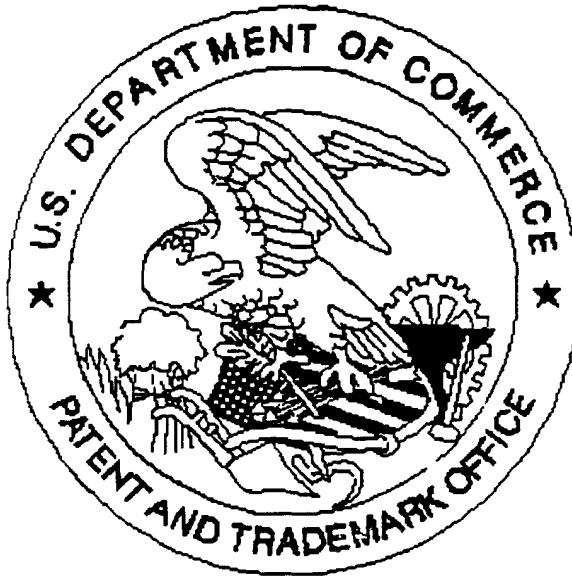
<212> DNA

<213> Escherichia coli

<400> 68

aacaaatctc agtcttctct gagctccgcc attgaacgtc tctcttctgg cctgcgtatt 60
 aacagtgcta aagatgacgc agcaggtcag gcgattgcta accgttttac agcaaatatt 120
 aaaggctga ctcaggcttc ccgtaacgcg aatgatggta tttctggtgc gcagaccact 180
 gaagggtgccc tttctgaaat caacaataac ttacagcgta ttogtgaatt gtcagtacag 240
 gccactaatg gtacaaactc tgactccgac ctgaattcaa ttcaggatga aattacacaa 300
 cgccttagtg aaattgatcg tgtttctaac cagacacaat ttaatgggtg aaaagttctg 360
 gcttctgac agactatgaa aattcaagta ggtgcgaacg atggtgaaac cattgagatt 420
 gcccttgata aaattgatgc taaaaccttg gggcttgata actttagcgt agcaccagga 480
 aaagtccaa tgtcctctgc ggttgcaact aagagcgaag ccgctcctga cttaactaag 540
 gtaaatgcaa ctgatggtag tgtgggaggt gctaaagcat tcggtagcaa ttataaaaaat 600
 gctgatgttg aaacttattt tgggtaccgg aatgtacaag atacaaggga tacaactgat 660
 gcgaccggta ctgcaggaaac aaaagtttat caagtacagg tgggaaggga gacttatttt 720
 gttgggtcaag ataataatac caacacgaac ggtttttacat tattgaaaca aaactctaca 780
 ggttatgaaa aagttcaggt ggggtggtgag gatgttcagt tagcaaaact tgggtggtcgt 840
 gtaactgcat ttgttgaaga taatgggtct gccacatcag ttgatttagc tgcgggtaaa 900
 atgggttaaag cattagctta taatgatgca ccaatgtctg tttatttttg gggaaaaaac 960
 ctagatgtcc accaagtaca agatacccaa gggaaatcctg tacctaattc atttgctgct 1020
 aaaacatcag accgcaccta cattgcagta aatgtagatg ccgctacagg taacacgtct 1080
 gttattactg atcctaattg taaggcagtt gaatgggcag taaaaaatga tggttctgca 1140
 caggcaatta tgcgtgaaga tgataagggt tatacagcca atatcacgaa taagacggca 1200
 accaaagggt ctgaactcag tgcctcagat ttgaaagcct tagcaaccac aaatccatta 1260
 tccacattag acgaagcttt ggcaaaaggt gataagttgc gcagttcttt ggggtgcagta 1320
 caaaaccgtt tcgactctgc catcaccaac cttggcaaca ccgtaaacaa cctgtcttct 1380
 gcccgtagcc gtatagaaga tgctgactac gcaaccgaag tgtctaacat gtctcgtgag 1440

United States Patent & Trademark Office
Office of Initial Patent Examination -- Scanning Division



SCANNED # 24

Application deficiencies found during scanning:

☒ Page(s) 63/96 of Drawings ^{was} ~~were~~ not present
for scanning. (Document title)

☒ Page(s) 7/7 of Sequences were not present
for scanning. (Document title)

☐ *Scanned copy is best available.*